

# Entanglement Spectrum Classification of $C_n$ -invariant Noninteracting Topological Insulators in Two Dimensions

Chen Fang<sup>1</sup>, Matthew J. Gilbert<sup>2,3</sup>, B. Andrei Bernevig<sup>1</sup>

<sup>1</sup>*Department of Physics, Princeton University, Princeton NJ 08544*

<sup>2</sup>*Department of Electrical and Computer Engineering,  
University of Illinois, Urbana IL 61801, USA and*

<sup>3</sup>*Micro and Nanotechnology Laboratory, University of Illinois, 208 N. Wright St, Urbana IL 61801, USA*

(Dated: August 24, 2012)

We study the single particle entanglement spectrum in 2D topological insulators which possess  $n$ -fold rotation symmetry. By defining a series of special choices of subsystems on which the entanglement is calculated, or real space cuts, we find that the number of protected in-gap states for each type of these real space cuts is a quantum number indexing (if any) non-trivial topology in these insulators. We explicitly show the number of protected in-gap states is determined by a  $Z^n$ -index,  $(z_1, \dots, z_n)$ , where  $z_m$  is the number of occupied states that transform according to  $m$ -th one-dimensional representation of the  $C_n$  point group. We find that the entanglement spectrum contains in-gap states pinned in an interval of entanglement eigenvalues  $[1/n, 1 - 1/n]$ . We determine the number of such in-gap states for an exhaustive variety of cuts, in terms of the  $Z_m$  quantum numbers. Furthermore, we show that in a homogeneous system, the  $Z^n$  index can be determined through an evaluation of the eigenvalues of point group symmetry operators at all high-symmetry points in the Brillouin zone. When disordered  $n$ -fold rotationally symmetric systems are considered, we find that the number of protected in-gap states is identical to that in the clean limit as long as the disorder preserves the underlying point group symmetry and does not close the bulk insulating gap.

The study of novel topological phases of matter has become one of the most active fields in condensed matter physics<sup>1–20</sup>. Chronologically speaking, the first of these phases to be experimentally realized is the integer quantum Hall (IQH) state, which is a striking departure from the traditional theory of conductivity due to its quantized Hall conductance and chiral edge modes<sup>21</sup>. Shortly after its experimental discovery, Thouless demonstrated that the special properties of the IQH state came down to its non-trivial bandstructure topology<sup>22</sup>. For the IQH state, one can prove that the quantized Hall conductance is equivalent to a momentum space integral of the Berry curvature, which is a non-zero integer (known as the Chern number in topology), multiplied by the conductance quantum  $e^2/h$ . Beyond this, Haldane demonstrated that the IQH state may be further generalized to a system which does not require an external overall magnetic flux, yet still possesses a non-zero Chern number and quantized Hall conductance. This system is generally referred to as a Chern insulator<sup>23</sup>.

Further as systems which preserve time-reversal invariance (TRI) pose an interesting problem as their Chern number vanishes in the presence of non-trivial topology. This necessitates the definition of a new quantum number capable of distinguishing between trivial and non-trivial topological states in systems which possess TRI. In 2D<sup>1–3</sup> and 3D<sup>7,24</sup>, one can define a  $Z_2$ -number, which here we call  $\gamma_0$ , as a quantity which is uniquely determined by the matrix representations of the time-reversal symmetry operator at all time-reversal invariant points within the Brillouin zone (BZ). In gapped systems,  $\gamma_0 = 1$  corresponds to topologically non-trivial insulators while  $\gamma_0 = 0$  corresponds to trivial insulators. In real mate-

rials, TRI topologically non-trivial insulators, or simply topological insulators usually have strong spin-orbit interactions which lead to an inversion of the bulk band gap at an odd number of time-reversal invariant points within the BZ<sup>15,25</sup>.

The important difference between Chern insulators (characterized by  $Z$ -indices) and TRI topological insulators (characterized by  $Z_2$ -indices) is the absence/presence of time-reversal symmetry. In general, we expect that the dimensionality and the symmetry of the system will determine the available quantum numbers that may be used to distinguish the topologically non-trivial states from the trivial ones. Motivated by this notion, Schnyder *et al.* classified translationally invariant topological insulators in presence/absence of three global symmetries: time-reversal, particle-hole, and chiral in 2D and 3D<sup>26,27</sup>. With translational symmetry, a non-interacting insulator maps to a manifold defined on the entire BZ. At each  $\mathbf{k}$  in BZ, the Hamiltonian maps to a projector onto the subspace spanned by the occupied Bloch states, or an element of the group  $U(N_{orb})/U(N_{occ}) \times U(N_{orb} - N_{occ})$ . The mathematical mappings of various insulators have been previously studied and it is found that they may be classified by their specific homotopy groups<sup>28</sup>. To be more concrete, in 2D the homotopy group for systems with only translational symmetry is  $\pi_2(U(N_{orb})/U(N_{occ}) \times U(N_{orb} - N_{occ})) = Z$ , which indicates that topological phases are indexed by integers. This picture is consistent with the previously understood conclusion that 2D systems may be characterized by a Chern number. In the same manner, TRI 2D and 3D insulators are also classified by the corresponding homotopy group,  $Z_2$ , of continuous mappings from

the  $d = 2, 3$ -dimensional BZ to a target space of projectors to occupied states under the constraint of TRI. Besides the bulk topological indices defined in insulators with translational invariance and with disorder, a bulk-edge correspondence exists which guarantees presence of gapless modes on the boundary between two insulators having different values of  $Z$  or  $Z_2$  index<sup>25,29,30</sup>.

The three symmetries considered above are local symmetries that do not involve spatial degrees of freedom. In condensed matter systems, we are faced with many types of symmetry operations in real space, besides time-reversal, charge conjugation and chiral symmetries. Amongst these symmetries, crystallographic point group symmetries (PGS) are well-known to universally exist within solids and may thus provide a natural avenue through which to continue the expansion of the family of materials which harbor topologically non-trivial states. Along this line, Fu<sup>31</sup> first studied spinless 3D insulators exhibiting gapless surface modes with quadratic band dispersion, the existence of which is protected by  $C_{4,6}$  rotation PGS and TRI. Furthermore, within the context of cold atoms, Sun *et al.* discussed a quadratic band crossing point in 2D that is protected by  $C_4$  PGS and TRI. These examples provide a tantalizing glimpse into the possibilities which exist for finding new non-trivial states of matter when PGS is combined with TRI. An equally interesting question one may ask is: can PGS bring about new topological phases, *without* the presence of TRI and how shall we identify them<sup>17,19</sup>? Without the presence of TRI, we are no longer *a priori* guaranteed to have any of the topological characteristics which are hallmarks of this symmetry such as: quantized magneto-electric coefficient or boundary/surface modes thereby necessitating the definition of a new topological quantum number.

To this end, the study of insulators with inversion symmetry<sup>32-34</sup>, which is the simplest yet non-trivial PGS, shows that in these insulators a new topological number may indeed be defined which is independent of the  $Z$ - or  $Z_2$  numbers previously introduced. This number can be used to characterize an inversion invariant insulator which contains a null result for both the Chern number ( $Z$ -number) and  $Z_2$ -number, but is still topologically distinct from a trivial insulator. The new quantum number labeling the new topological insulators is the number of protected in-gap states in the single particle entanglement spectrum.

The entanglement spectrum has been successfully applied to identify topological orders in a variety of condensed matter systems such as fractional quantum Hall systems<sup>35</sup>, spin chains<sup>36</sup>, Chern insulators<sup>37</sup>, and topological band insulators<sup>30,38-40</sup>. Unlike the energy spectrum, the quantum entanglement spectrum solely depends on the many-body ground state, and is, for this reason, more suitable for identifying any non-trivial topology in the ground state. In noninteracting systems, it can be shown that to obtain the many-body entanglement spectrum, only the single particle entanglement spectrum is needed, defined as all eigenvalues of the one-

body reduced density matrix<sup>41,42</sup>:

$$C_{ij}(A) = \langle c_i^\dagger c_j \rangle, \quad (1)$$

where  $i, j \in A$  and  $A$  is a sub-system of the whole system  $L$ . In principle, all ‘energies’ in the entanglement spectrum are between zero and unity corresponding to states which are either localized outside of or inside the sub-system, respectively. In the case of inversion symmetric topological insulators<sup>32</sup>, if  $A$  is chosen as the left/right half of the  $L$ , then there can be several eigenstates of  $C(A)$  with degenerate eigenvalue at exact  $1/2$ , in an inversion invariant insulator. These degenerate eigenvalues correspond to states that are topologically protected in the sense that they always stay at pinned  $1/2$  so long as the inversion symmetry is preserved and the bulk gap remains open. These states are referred to as the protected  $1/2$ -in-gap states and their number defines a new topological number to be used to differentiate between trivial and non-trivial inversion invariant insulators.

In this work, we focus on insulators with another type of PGS,  $n$ -fold rotation symmetry where  $n = 2, 3, 4, 6$  by lattice restriction. We show that for proper choices of sub-system  $A$  in PGS invariant insulators, there exist protected in-gap states the number of which defines a new topological invariant. While these states are analogous to the  $1/2$ -in-gap states previously seen in inversion invariant insulators, two key differences are noted: first, instead of taking  $A$  as the left/right half of  $L$ , there exists more than one type of symmetric cuts for a  $C_n$  invariant insulator and, in general, each has its own corresponding number of in-gap states. Second, the protected in-gap states now are not necessarily located at exact  $1/2$ , but within a region around  $1/2$ , prevented by symmetry from arbitrarily approaching two ends of the single particle entanglement spectrum (zero and unity). We calculate the number of protected in-gap states for each type of symmetric cuts in a  $C_n$  invariant insulator in terms of  $m$ -fold rotation eigenvalues at high-symmetry points in the BZ, where  $m$  is a factor of  $n$ . Further, we show that all the numbers of protected in-gap states can be expressed in terms of a  $Z^n$ -index ( $z_1, z_2, \dots, z_n$ ), giving a  $Z^n$ -classification of these insulators. The physical reason underlying the  $Z^n$  classification is very simple: if the single particle Hamiltonian is  $n$ -fold invariant, then every single particle wavefunction belongs to a certain 1D representation of  $C_n$  group. The number of electrons in each representation below the Fermi level then naturally gives the index ( $z_1, z_2, \dots, z_n$ ). Finally, we will show that if the translational symmetry is broken by introducing disorder potential that is also  $C_n$  invariant, the  $Z^n$ -index remains unchanged and as do the associated numbers of protected in-gap states.

The paper is organized as follows. In Sec.I, we introduce concepts related to our examination of PGS insulators in 2D referred to throughout the paper: In Sec.I(A), we introduce crystallographic point groups; in Sec.IB, we discuss a general definition of non-interacting topologically non-trivial insulator, and in Sec.I(C), a brief in-

roduction to single particle entanglement spectrum and its relation to entanglement entropy. In Sec.II we revisit the entanglement spectrum in a 1D inversion invariant system. In this revisit, we analytically prove the relation between the number of protected in-gap states and the inversion eigenvalues at  $k = 0$  and  $k = \pi$ . The emphasis here is the relation between the number of 1/2-in-gap states and a  $Z^2$ -index  $(z_1, z_2)$ , where  $z_1$  ( $z_2$ ) is the number of occupied Bloch states with odd (even) parity. Sec.III we perform a simple extension of the previous 1D results to 2D and 3D inversion invariant insulators, where we relate the number of 1/2-in-gap states to the parity of Chern number and, when considering systems with TRI, to the  $Z_2$  number. In Sec.IV, we present our main result. We outline the analytic proof by which we relate the  $Z^n$  index to the number of protected in-gap states in a  $C_n$  invariant insulator with every type of possible symmetric cut, with some specific details relegated to the various Appendices. In Sec.V, we briefly discuss 3D PGS invariant insulators in Sec.V(A); we discuss in Sec.V(B) the effect of weak interaction, which reduces the  $Z^n$ -index down to a  $Z_n$ -index as single particle states lose their meaning while the *many-body* wavefunction of any non-degenerate insulating state is still a 1D representation and can be associated with some  $Z_n$  number.

## I. PRELIMINARIES

Before proceeding with the relevant background material, we briefly clarify the notations we use throughout the paper. When we use the notation  $C_n$ , we are referring to either the point group  $C_n$ , or to the  $n$ -fold rotation, the choice of which should be clear given the context.  $\hat{C}_n$  refers an operator in the Hilbert space corresponding to the  $n$ -fold rotation.  $\mathcal{C}_n$  denotes the matrix representation of the operator  $\hat{C}_n$  in a basis, either the orbital basis or the occupied-band basis, which should be discernible given the context of its usage. Similar definitions and distinctions apply to other operators, e.g.,  $M_z$ ,  $\hat{M}_z$ ,  $\mathcal{M}_z$  refer to mirror symmetry  $z \rightarrow -z$ , its operator in Hilbert space, and a matrix representation of the operator, respectively. Finally, when used as indices, the Greek letters will always span the orbitals, while Roman letters denote either the bands or spatial directions ( $x, y, z$  or  $k_x, k_y, k_z$ ). In all equations, repeated indices are, unless otherwise noted, automatically summed for the following types of summations: summation over all sites in the system, summation over all sites in the subsystem on which the entanglement spectrum is calculated, summation over all orbitals, summation over all bands (occupied and unoccupied).

### A. 2D Crystallographic Point Groups

A point group is a group of all symmetry operations (an atom, a molecule or a lattice) that leave at least one

point fixed in space. In nature, there exist an infinite number of point groups. On the other hand, crystallographic point groups are point groups that are consistent with lattice translational symmetry. The crystallographic restriction theorem states that lattice translational symmetry dictates that  $n = 1, 2, 3, 4, 6$  for any  $n$ -fold rotation axis and effectively limits the total number of crystallographic point groups. In 2D there are 10 different point groups allowed which are comprised of five cyclic groups and five dihedral groups. A cyclic point group  $C_n$  consists of all powers of an  $n$ -fold rotation while a dihedral point group  $D_n$  contains  $C_n$  as a subgroup and an additional two-fold in-plane rotation axis. In this paper we will focus on  $C_n$  PGS.

With the point groups defined, we now consider the representations of point groups in a single particle Hilbert space. In the representation, every symmetry operation corresponds to an operator in the Hilbert space. As we will discuss both *spinless* and *spinful* fermions, it is necessary at this point to introduce double point groups,  $G^D$ . Mathematically speaking,  $G^D = \{E, \bar{E}\} \otimes G$ , where  $E$  and  $\bar{E}$  are the identity operator and a rotation of angle  $2\pi$  about any axis in space, which satisfies  $\bar{E}^2 = E$ , and  $\bar{E}g = g\bar{E}$  for  $g \in G$ . This enlarged group has physical meanings as one considers its representations, which are classified into two types: in the first type  $D(E) = D(\bar{E}) = \mathcal{I}$ , where  $\mathcal{I}$  is the identity matrix; in the second type  $D(E) = -D(\bar{E}) = \mathcal{I}$ . Representations of the first type are referred to as single-valued and representations of the second type are called double-valued. A degenerate subspace, in which every state has the same energy, of a single fermion must form a single-valued/double-valued representation of the underlying PGS of the Hamiltonian if the particles are spinless/spinful, as we know that a spinful fermion takes an extra minus sign as it is rotated by  $2\pi$ . For  $C_n$  PGS, as one applies the general statement to Hilbert space representation of  $C_n$  PGS, one can obtain  $\hat{C}_n^n = (-1)^F \hat{I}$ , and if in the orbital/band space representations, similarly,  $\mathcal{C}_n^n = (-1)^F \mathcal{I}$ , where  $F = 0$  ( $F = 1$ ) for spinless (spinful) fermions.

Given the Hilbert space representation  $\hat{R}$  of a point group operator  $R$ , we now look at the sufficient and necessary conditions for a single particle tight-binding Hamiltonian  $\hat{H}$  to be invariant under this operator. We assume that in one unit cell at  $\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$  there are  $s$  orbitals at  $\mathbf{r} = \mathbf{R} + \mathbf{d}$ , where  $\mathbf{d}$  is the offset vector of the atom from the nearest lattice point. Note that we have ignored all unit cell structures, a simplification that spares us from discussing non-symmorphic space groups, which contain operations combined from member of a point group and translations by fractions of lattice vectors. Now the PGS operator  $R$  sends an electron at  $\mathbf{r}$  to  $R\mathbf{r} = R\mathbf{R} + R\mathbf{d}$ . If  $R$  is a symmetry, there must be another atom at  $R\mathbf{r}$ , i.e.,  $\mathbf{R}' = R\mathbf{R} + R\mathbf{d} - \mathbf{d}$  is also a lattice point. Writing the above relation in second quantization,

we have

$$\hat{R}c_\alpha(\mathbf{R})\hat{R}^{-1} = \sum_{\beta} \mathcal{R}_{\alpha\beta}c_\beta(\mathbf{R}'). \quad (2)$$

In this equation,  $\mathcal{R}_{\alpha\beta}$  describes a possible rotation of an orbital under  $R$ . For example, suppose we are considering  $p_{x,y,z}$ -orbitals under  $R = C_4$  a four-fold rotation about  $z$ -axis, then obviously,  $p_x$  is sent to  $p_y$ ,  $p_y$  to  $-p_x$  and  $p_z$

to  $p_z$ . In this case we have in the basis of  $(p_x, p_y, p_z)$

$$\mathcal{R}_4 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3)$$

If we adopt the following definitions of Fourier transform of  $c_\alpha(\mathbf{R} + \mathbf{d})$

$$c_\alpha(\mathbf{k}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} c_\alpha(\mathbf{R}) \exp(-i\mathbf{k} \cdot \mathbf{R}), \quad (4)$$

from Eq.(2), we obtain:

$$\begin{aligned} \hat{R}c_\alpha(\mathbf{k})\hat{R}^{-1} &= \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} \mathcal{R}_{\alpha\beta}c_\beta(\mathbf{R}') \exp(-i\mathbf{k} \cdot \mathbf{R}) \\ &= \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} \mathcal{R}_{\alpha\beta}c_\beta(\mathbf{R}') \exp[-i\mathbf{k} \cdot (R^{-1}\mathbf{R}' + R^{-1}\mathbf{d} - \mathbf{d})] \\ &= \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} \mathcal{R}_{\alpha\beta}c_\beta(\mathbf{R}') \exp[-i(R\mathbf{k}) \cdot (\mathbf{R}' + \mathbf{d} - R\mathbf{d})] \\ &= \mathcal{R}_{\alpha\beta}c_\beta(R\mathbf{k})e^{i\mathbf{k} \cdot (\mathbf{d} - R\mathbf{d})} \\ &\equiv \bar{\mathcal{R}}_{\alpha\beta}(\mathbf{k})c_\beta(R\mathbf{k}). \end{aligned} \quad (5)$$

It is easy to verify that  $\bar{\mathcal{R}}(\mathbf{k})$  is a unitary matrix using that  $\hat{R}$  is a unitary symmetry operation. A tight-binding model Hamiltonian with translational symmetry  $\hat{H} = \sum_{\mathbf{k}} \mathcal{H}_{\alpha\beta}(\mathbf{k})c_\alpha^\dagger(\mathbf{k})c_\beta(\mathbf{k})$  transforms under  $\hat{R}$  as

$$\hat{R}\hat{H}\hat{R}^{-1} = \sum_{\mathbf{k}} (\bar{\mathcal{R}}(\mathbf{k})\mathcal{H}(\mathbf{k})\bar{\mathcal{R}}^{-1}(\mathbf{k}))_{\alpha\beta}c_\alpha(R\mathbf{k})c_\beta(R\mathbf{k}) \quad (6)$$

If  $R$  is a symmetry, we have  $\hat{R}\hat{H}\hat{R}^{-1} = \hat{H}$  therefore

$$\bar{\mathcal{R}}(\mathbf{k})\mathcal{H}(\mathbf{k})\bar{\mathcal{R}}^{-1}(\mathbf{k}) = \mathcal{H}(R\mathbf{k}). \quad (7)$$

Using  $\bar{\mathcal{R}}(\mathbf{k}) = \mathcal{R}e^{i\mathbf{k} \cdot (\mathbf{d} - R\mathbf{d})}$ , we have

$$\mathcal{R}\mathcal{H}(\mathbf{k})\mathcal{R}^{-1} = \mathcal{H}(R\mathbf{k}). \quad (8)$$

Eq.(8) determines the sufficient and necessary conditions for a single-particle tight-binding Hamiltonian to be invariant under a point group symmetry  $R$ . For a  $C_n$  invariant 2D system, the above equation simplifies to

$$\begin{aligned} \mathcal{C}_n\mathcal{H}(k_x, k_y)\mathcal{C}_n^{-1} &= \\ \mathcal{H}(k_x \cos(\frac{2\pi}{n}) - k_y \sin(\frac{2\pi}{n}), k_x \sin(\frac{2\pi}{n}) + k_y \cos(\frac{2\pi}{n})), \end{aligned} \quad (9)$$

where  $\mathcal{C}_n$  is the transformation matrix in the orbital basis for the  $n$ -fold rotation about  $z$ -axis.

## B. Generalized definition of a topological insulator

In this paper, we define a topological insulator as a system the many-body ground state of which cannot be adiabatically deformed into the ground state in the atomic

limit<sup>32,43</sup>. The atomic limit is the limit of zero coupling between orbitals on different atoms but with all symmetries of the system preserved. Every system in this limit is considered as a trivial insulator. Clearly then, insulators that cannot be continuously tuned to the atomic limit while preserving the gap and all symmetries are nontrivial topological insulators according to this definition. This definition enables us to use any quantized quantity depending on the wavefunction to differentiate topologically trivial and non-trivial insulators. This is because a quantized quantity only changes when gap is closed, thus if the value of this quantity in an insulator is different than that in its atomic limit, it must be a topologically non-trivial insulator, or simply, a topological insulator. Such a quantity is called a topological index, the choice of which is determined completely by the dimensionality and the given symmetries of a Hamiltonian. This definition comprises more types of non-trivial topologies than those that may be labeled by a non-zero  $Z$  or  $Z_2$  index. Furthermore, it should be noted that a good quantum number that denotes a topological state is not necessarily a physical observable such as Hall conductance. Such a case has been noted previously in non-trivial topological insulators possessing inversion symmetry. These systems have topologically protected in-gap states at exactly 1/2 in the entanglement spectrum and thus they cannot be removed by adiabatic deformations of the Hamiltonian<sup>32</sup>. Nevertheless, while the number of in-gap states is a topological index, it is not related to any yet known experimental observable.



### C. Single Particle Entanglement Spectrum

We have defined the general single particle entanglement spectrum in Eq.(1), and here we restrict the definition to the entanglement spectrum associated with a special type of subsystem: *real space* cuts. To begin, the correlation matrix for a real space cut is given by

$$C_{i\alpha,j\beta} = \langle c_{\alpha}^{\dagger}(\mathbf{r}(i))c_{\beta}(\mathbf{r}(j)) \rangle, \quad (10)$$

where  $\mathbf{r}(i)$  denotes a unit cell within a predefined subset of the system, called  $A$ , and  $\alpha, \beta$  denote orbitals in a unit cell. Suppose in the selected subset one has  $N_A$  unit cells and each unit cell has  $N_{orb}$  orbitals, the entanglement is then given by a  $N_A * N_{orb} \times N_A * N_{orb}$  matrix, where  $N_A$  is the number of sites in the real space cut  $A$ . The set of eigenvalues of the matrix is defined as the single-particle entanglement spectrum or simply entanglement spectrum associated with subset  $A$ . The definition is a special one as it confines the  $c$ -operators within a predefined set of sites in real space. In general, there are other types of cuts such as momentum cuts or orbital cuts<sup>35</sup>, which confine the  $c$ -operators within a range of momentum or a set of orbitals (e.g., Landau orbitals). From definition, it is easy to show that all eigenvalues of  $C$ , i.e., all ‘energies’ in the spectrum, are within the range  $[0, 1]$ . A cut may be chosen as the whole system,  $L$ , and one can prove that in this case the correlation matrix  $C(L)$  must be a projector matrix, whose eigenvalue is either zero or unity<sup>32</sup>.

The reason we are only interested in real space cuts in this paper is that we will focus on band insulators, whose constituent single particle wavefunctions are completely local in momentum space with  $\mathbf{k}$  as a good quantum number. Therefore all eigenvalues in a momentum cut must be either zero or unity, a situation usually referred to as trivial. On the other hand, a real space cut has a potential to distinguish insulators that can be adiabatically tuned into the atomic limit and those that cannot be. To better understand this point, let us consider an insulator in the atomic limit. In this limit all single particle wavefunctions are localized, and therefore when a real space cut is made, a particle must be either totally inside or totally outside the subsystem defined by the cut. This forces the eigenvalue of the correlation matrix to be either 1 if the particle is located within the cut or 0 if the particle is outside of the cut. If an insulator *cannot* be adiabatically tuned to the atomic limit, *then within the entanglement spectrum there must be eigenvalues that are topologically prevented from approaching arbitrary vicinities of 1 or 0*. We may gain additional insight into the entanglement spectrum by considering a 2D IQH state, or a Chern insulator, that has  $C$  (=Chern number) protected extended states and, when a real space cut is made, there must be at least the same number of states that are neither inside nor outside the subsystem defined by the cut. These extended topological states present themselves as protected ‘in-gap’, neither 0 nor 1, states in the entanglement spectrum. This means that existence of protected

in-gap states in the entanglement spectrum implies existence of intrinsically delocalized states.

The single particle entanglement spectrum is closely related to the entanglement entropy<sup>42,44–46</sup> by

$$E_q = - \sum_n [\xi_n \ln \xi_n + (1 - \xi_n) \ln(1 - \xi_n)], \quad (11)$$

where  $\xi_n$  is the  $n$ th eigenvalue in the spectrum. Entanglement entropy is known to quantify non-trivial topology in insulators: If the entanglement entropy cannot be adiabatically tuned to zero, the insulator is topologically non-trivial. When there are protected in-gap states in the entanglement spectrum, then from Eq.(11) the entanglement entropy always remain finite. Therefore, presence of protected in-gap states indicates non-trivial topology in an insulator.

## II. ENTANGLEMENT SPECTRUM OF INVERSION INVARIANT TOPOLOGICAL INSULATORS IN 1D: A REVISIT

### A. Decomposition of the correlation matrix into projectors

Let us begin our study of the entanglement spectrum of point group symmetric topological insulators by examining the inversion invariant system<sup>32,34</sup> in 1D. The system is assumed to have  $N$  sites where  $N$  is an *even* number. The inversion center is *at the midway* of the  $N/2$ -th site and  $(N/2 + 1)$ -th site. For now, we assume that the system is translationally invariant, hence from Eq.(8) we have  $[\mathcal{P}, \mathcal{H}(0)] = [\mathcal{P}, \mathcal{H}(\pi)] = 0$ , by which we can label all eigenstates at 0 and  $\pi$  with their eigenvalues under  $\mathcal{P}$  ( $\pm 1$ ).  $k = 0$  and  $k = \pi$  are called the invariant  $k$ -points under inversion, denoted by  $k_{inv}$ . For a reason that is detailed in Appendix A, we now use, instead of eigenvalues of  $\mathcal{P}$ , eigenvalues of  $\bar{\mathcal{P}}(k) = \mathcal{P}e^{ik(d-Pd)}$  to label the eigenstates at  $k = 0, \pi$ . For inversion invariant system, the symmetry fixed point is in the middle of two sites so  $d = 1/2$ , and  $d - Pd = 1/2 - (-1/2) = 1$ . From this we have  $\bar{\mathcal{P}}(k) = e^{ik}\mathcal{P}$ .

Now we separate the left-half of the system, from the 1st to the  $N/2$ th site, as the subsystem  $A$  and calculate the entanglement spectrum  $\{\xi_n|A\}$ . There can be eigenvalues in the spectrum lying exactly at  $\xi_n = 1/2$ , and the number of these eigenvalues is given by

$$N_{1/2} = |\sum_{k_{inv}} (n^+(k_{inv}) - n^-(k_{inv}))|, \quad (12)$$

where  $n^+(k_{inv})$  and  $n^-(k_{inv})$  represent the number of *occupied* states with  $+1$  and  $-1$  eigenvalue of  $\bar{\mathcal{P}}(k_{inv})$  of the  $m$ th band, respectively. If there is an impurity potential, it will generally split the  $1/2$ -in-gap degeneracy; however, if the impurity potential is also inversion symmetric, having equal magnitude at the  $i$ -th site and the  $(N + 1 - i)$ -th site, the mid-gap states are preserved.

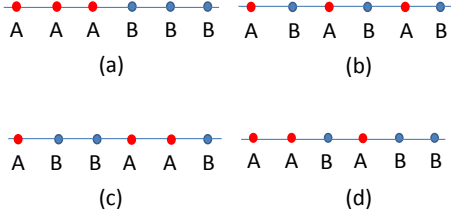


FIG. 1. Examples of inversion symmetric cut in a 1D system with six sites. All sites in red define the real space cut on which the entanglement spectrum is calculated.

In fact, we desire to make the statement more general by showing that the same result holds for all ‘inversion symmetric cuts’, of which the above left-half cut is a special case. The subset which defines an inversion symmetric cut does not have to be the left or the right half, but should satisfy the following conditions: First, the subset  $A$  has exactly  $N/2$  sites, and second, if site  $i$  is in  $A$ , site  $N+1-i$  must not be in  $A$ . Such a subset  $A$  will be referred to as an inversion symmetric cut. A set containing all sites left to the inversion center is a symmetric cut, and several other possible symmetric cuts are plotted in Fig.1. Below, we will use  $r_i$  to denote the position of the  $i$ -th site in the cut for  $i = 1, \dots, N/2$ . When the sum is over  $i$ , it means summing over  $r_i$ ; when the sum is over  $r$ , it means summing over all sites in  $L$ .

While the correlation matrix  $C$  in Eq.(10) is not a projector, if  $A$  is an inversion symmetric cut in an inversion invariant insulator,  $C$  can be expressed as the linear average of two projectors  $D$  and  $\bar{D}$ . To show this, we define new operators:

$$\begin{aligned} d_{i\alpha} &= (c_\alpha(r_i) + \hat{P}c_\alpha(r_i)\hat{P})/\sqrt{2}, \\ \bar{d}_{i\alpha} &= (c_\alpha(r_i) - \hat{P}c_\alpha(r_i)\hat{P})/\sqrt{2}, \end{aligned} \quad (13)$$

where  $i = 1, \dots, N/2$  and  $\alpha$  is the orbital index (including spin). These are the symmetric and antisymmetric combinations of the annihilation operators. The single particle entanglement matrix for subsystem  $A$  is:

$$\begin{aligned} C_{i\alpha,j\beta} &= \langle c_\alpha^\dagger(r_i)c_\beta(r_j) \rangle \\ &= \frac{1}{2} \langle (d_{i\alpha}^\dagger + \bar{d}_{i\alpha}^\dagger)(d_{j\beta} + \bar{d}_{j\beta}) \rangle. \end{aligned} \quad (14)$$

Then we point out a simple fact:

$$\langle d_{i\alpha}^\dagger \bar{d}_{j\beta} \rangle = \langle \bar{d}_{i\alpha}^\dagger d_{j\beta} \rangle = 0. \quad (15)$$

To prove this one only needs to notice that the ground state wavefunction must be an eigenstate of  $\hat{P}$ , and that  $\hat{P}d_{i\alpha}\hat{P} = d_{i\alpha}$  while  $\hat{P}\bar{d}_{i\alpha}\hat{P} = -\bar{d}_{i\alpha}$ . Using Eq.(15), we have

$$C_{i\alpha,j\beta} = \frac{1}{2}(D_{i\alpha,j\beta} + \bar{D}_{i\alpha,j\beta}), \quad (16)$$

where

$$\begin{aligned} D_{i\alpha,j\beta} &= \langle d_{i\alpha}^\dagger d_{j\beta} \rangle, \\ \bar{D}_{i\alpha,j\beta} &= \langle \bar{d}_{i\alpha}^\dagger \bar{d}_{j\beta} \rangle. \end{aligned} \quad (17)$$

Next we prove that  $D$  and  $\bar{D}$  are projector matrices, i.e.,  $D^2 = D$  and  $\bar{D}^2 = \bar{D}$ . In an inversion invariant system, all single particle eigenstates can be divided into two subspaces with even and odd parity respectively. We use  $\psi_n$  to denote the annihilation operator of an even parity eigenstate and  $\bar{\psi}_n$  to denote that of an odd parity eigenstate, i.e.,  $\hat{P}\psi_n\hat{P} = \psi_n$  and  $\hat{P}\bar{\psi}_n\hat{P} = -\bar{\psi}_n$ . Here  $n$  is not a band label as before, because we do not have translational symmetry, but simply an index counting all single particle states. In a system with  $N$  sites and  $N_{orb}$  orbitals per site, there are  $NN_{orb}/2$  states with even and odd parity respectively, so  $n$  ranges from 1 to  $NN_{orb}/2$  in general. Every creation/annihilation operator can be expressed in this basis and we have the unitary transform

$$c_{r\alpha} = U_{r\alpha,n}\psi_n + V_{r\alpha,n}\bar{\psi}_n. \quad (18)$$

Then from Eq.(13) we have  $d_{r\alpha} = \sqrt{2}U_{r\alpha,n}\psi_n$ . On the other hand we have

$$\psi_n = U_{n,r\alpha}^\dagger c_{r\alpha}, \quad (19)$$

$$\begin{aligned} \psi_n &= \hat{P}\psi_n\hat{P} \\ &= U_{n,r\alpha}^\dagger \hat{P}c_{r\alpha}\hat{P} \\ &= U_{n,r\beta}^\dagger \mathcal{P}_{\beta\alpha} c_{N+1-r\alpha}. \end{aligned} \quad (20)$$

Equating right hand sides of the last two equations, we have

$$U_{r\alpha,n} = \mathcal{P}_{\alpha\beta} U_{N+1-r\beta,n}. \quad (21)$$

Using these relations we can prove that  $D$  is a projector by calculating  $D^2$  directly. In fact,

$$D_{j\beta,i\alpha} = \langle d_{j\beta}^\dagger d_{i\alpha} \rangle = 2 \sum_{m \in occ.} U_{m,r_j\beta}^\dagger U_{r_i\alpha,m}. \quad (22)$$

And we have

$$D_{j\beta,i\alpha} D_{l\gamma,j\beta} = 4 \sum_{m,n \in occ.} U_{r_i\alpha,m} U_{m,r_j\beta}^\dagger U_{r_j\beta,n} U_{n,r_l\gamma}^\dagger, \quad (23)$$

where the summation is over the sites belonging to the  $A$ -subsystem.

Care should be taken to note that  $U_{m,r_j\beta}^\dagger U_{r_j\beta,n} \neq \delta_{mn}$ ! This is because by summing  $j$  we only sum half of the system. When inserting Eq.(21) we obtain:

$$\begin{aligned} U_{m,r_j\beta}^\dagger U_{r_j\beta,n} &= U_{m,N+1-r_j\alpha}^\dagger \mathcal{P}_{\alpha\beta}^\dagger \mathcal{P}_{\beta\gamma} U_{N+1-r_j\gamma,n} \\ &= U_{m,N+1-r_j\beta}^\dagger U_{N+1-r_j\beta,n} \text{ [Do not sum over } j]. \end{aligned} \quad (24)$$

This leads to

$$U_{m,r_j\beta}^\dagger U_{r_j\beta,n} = \frac{1}{2} U_{m,r\beta}^\dagger U_{r\beta,n} = \frac{\delta_{mn}}{2}. \quad (25)$$

Substituting Eq.(25) into Eq.(23), we see  $D^2 = D$ , or  $D$  is a projector.

Using Eqs.(22,25), it is easy to see that every row in  $U_{r_j\alpha,n}$  is an eigenvector of  $D$ . If  $\psi_n$  annihilates an occupied state, this vector has eigenvalue 1 and if  $\psi_n$  annihilates an unoccupied state, it has eigenvalue 0. So

the dimension of  $D$ , or the number of unity eigenvalues is:  $\dim(D) = \text{Dim}(\Psi_{occ.})$ , or the number of occupied eigenstates with even parity. Similarly, it can be shown that  $\dim(\bar{D}) = \text{Dim}(\bar{\Psi}_{occ.})$ , or the number of occupied eigenstates with odd parity.

If  $\dim(D) > \dim(\bar{D})$ , there must be at least  $\dim(D) - \dim(\bar{D})$  common eigenstates of  $D$  and  $\bar{D}$  with unity and zero eigenvalues, respectively. These states are therefore eigenstates of  $C$  with eigenvalue exactly at  $1/2$ , i.e.,  $1/2$ -in-gap states. And if  $\dim(D) < \dim(\bar{D})$ , there must be  $\dim(\bar{D}) - \dim(D)$   $1/2$ -in-gap states. In general, we have  $N_{1/2} = |\dim(D) - \dim(\bar{D})|$ . The two integers  $\dim(D) = \text{Dim}(\Psi_{occ})$  and  $\dim(\bar{D}) = \text{Dim}(\bar{\Psi}_{occ})$  are the numbers of occupied eigenstates with even and an equal number of states with odd parities, or in other words, they are eigenstates in the two 1D representations of the inversion point group. These two numbers,  $(z_1, z_2) = (\text{Dim}(\Psi_{occ}), \text{Dim}(\bar{\Psi}_{occ}))$ , give a  $Z^2$  classification of inversion invariant insulators. The sum of its two components  $z_1 + z_2$  is simply the total number of fermions, and their difference  $|z_1 - z_2|$  is exactly the number of protected in-gap states in the entanglement spectrum, i.e.,

$$N_{1/2} = |z_1 - z_2|. \quad (26)$$

In a trivial insulator, one has  $z_1 = z_2$ , while any pair  $(z_1, z_2)$  with  $z_1 \neq z_2$  denotes a topologically non-trivial inversion invariant insulator. Moreover, it should be noted that the proof presented here, up to this point, does not require translational symmetry.

### B. Translational invariance and number of in-gap states

Now we may proceed to calculate  $\text{Dim}(\Psi_{occ})$  and  $\text{Dim}(\bar{\Psi}_{occ})$  by relating them to the inversion eigenvalues at high symmetry  $k$ -points  $0$  and  $\pi$  in presence of translational symmetry. Suppose  $\gamma_m(k)$  is the annihilation operator of the Bloch wavefunction at  $k$  on the  $m$ th band, and define symmetry adapted operators

$$\begin{aligned} \psi_m(k) &= (\gamma_m(k) + \hat{P}\gamma_m(k)\hat{P})/\sqrt{2}, \\ \bar{\psi}_m(k) &= (\gamma_m(k) - \hat{P}\gamma_m(k)\hat{P})/\sqrt{2}, \end{aligned} \quad (27)$$

and they satisfy  $\hat{P}\psi_m(k)\hat{P} = \psi_m(k)$ ,  $\hat{P}\bar{\psi}_m(k)\hat{P} = -\bar{\psi}_m(k)$ . Due to inversion symmetry, the single particle states  $|\psi(k)\rangle = \psi^\dagger(k)|0\rangle$  and  $|\bar{\psi}(k)\rangle = \bar{\psi}^\dagger(k)|0\rangle$ , in which  $|0\rangle$  refers to the vacuum state and not the Fermi sea, are the eigenstates of  $\hat{H}$ . The two sets together, for  $m \in occ$ , make a basis of the filled bands in which every basic vector has a certain parity. The subspaces  $\Psi_{occ}^{(+1)}$  and  $\Psi_{occ}^{(-1)}$  are given by:  $\Psi_{occ} = \{|\psi_m(k)\rangle | m \in occ., k \geq 0, \psi_m(k) \neq 0\}$  and  $\bar{\Psi}_{occ} = \{|\bar{\psi}_m(k)\rangle | m \in occ., k \geq 0, \bar{\psi}_m(k) \neq 0\}$ . We observe that we have a rather strange looking constraint  $\psi_m(k) \neq 0$  and  $\bar{\psi}_m(k) \neq 0$ . This is because at  $k = 0, \pi$ , either  $\psi_m(k_{inv}) = 0$  or  $\bar{\psi}_m(k_{inv}) = 0$ , and either the state  $|\psi_m(k)\rangle$  or  $|\bar{\psi}_m(k)\rangle$  becomes a null state.

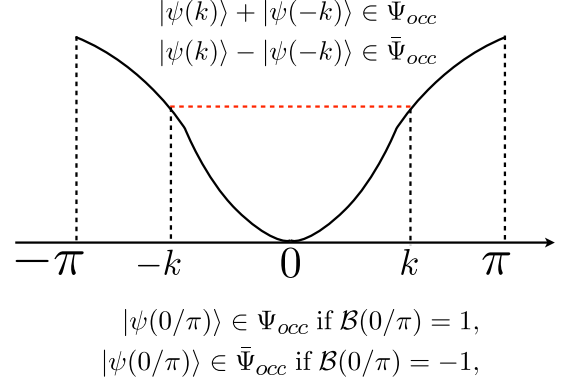


FIG. 2. A schematic of a single band in a 1D inversion symmetric insulator. From the figure we clearly see that at any  $k = 0$  we have two equal energy states at  $k$  and  $-k$ , while at  $k = 0, \pi$ , there is only one state. The two degenerate states at  $\pm k$  can be recombined into two states with even and odd parities respectively, while the singlet state at  $0$  or  $\pi$  is either even or odd.

To see this, we introduce the sewing matrix

$$\mathcal{B}_{mn}(k) = \langle u_m(-k) | \hat{P} | u_n(k) \rangle \quad (m, n \in occ). \quad (28)$$

At  $k = 0$  and  $k = \pi$ , Hamiltonian commutes with inversion, and, therefore, the sewing matrix must be diagonal with diagonal elements being either  $+1$  or  $-1$ . Applying the sewing matrix by utilizing a property proven in Appendix B, we have

$$\hat{P}\gamma_m(k)\hat{P} = \mathcal{B}_{nm}\gamma_n(-k). \quad (29)$$

Inserting the above equation into Eq.(27) and take  $k = 0, \pi$ , we obtain

$$\begin{aligned} \psi_m(k_{inv}) &= (1 + \mathcal{B}_{mm}(k_{inv}))\gamma_m(k_{inv})/\sqrt{2}, \\ \bar{\psi}_m(k_{inv}) &= (1 - \mathcal{B}_{mm}(k_{inv}))\gamma_m(k_{inv})/\sqrt{2}. \end{aligned} \quad (30)$$

It is then obvious that if  $\mathcal{B}_{mm}(k_{inv}) = 1$ , then  $\bar{\psi}_m(k_{inv}) = 0$ , otherwise  $\psi_m(k_{inv}) = 0$ . This is the mathematical description of the rather simple fact that for each at  $k \neq 0, \pi$ , there is its counterpart at  $-k$ , from which we can construct two states with even and odd parities respectively, while at  $k = 0, \pi$  we cannot do this and either the even or the odd state must be a null state. The counting of states in  $\Psi_{occ}$  and  $\bar{\Psi}_{occ}$  can be done on each band separately and then we add the numbers together. At  $k \neq 0, \pi$ ,  $\psi_m(k)$  and  $\bar{\psi}_m(k)$  appear in pairs so they contribute equally to  $\text{Dim}(\Psi_{occ})$  and  $\text{Dim}(\bar{\Psi}_{occ})$  (see Fig.2 for an example). The number of states in  $\Psi_{occ}$  and  $\bar{\Psi}_{occ}$  contributed by this part is  $(N - 2)/2$  for each. Counting the states at inversion invariant points needs information of the sewing matrix.

To be more concrete, let us examine a system with only one band filled. Suppose we have  $B(0) = B(\pi) = 1$ ,

using Eq.(30), then

$$\begin{aligned}\psi(0) &= \sqrt{2}\gamma(0), \\ \bar{\psi}(0) &= 0, \\ \psi(\pi) &= \sqrt{2}\gamma(\pi), \\ \bar{\psi}(\pi) &= 0.\end{aligned}\tag{31}$$

Both at  $k = 0$  and  $k = \pi$ , there is one more state in  $\Psi_{occ}$  than in  $\bar{\Psi}_{occ}$ . From this we see that  $z_1 = \text{Dim}(\Psi_{occ}) = (N - 2)/2 + 2$  and  $z_2 = \text{Dim}(\bar{\Psi}_{occ}) = (N - 2)/2$ . According to previous discussion, we have  $|z_1 - z_2| = 2$  1/2-in-gap states. One can repeat the discussion for  $B(0) = B(\pi) = -1$  and find  $z_1 = \text{Dim}(\Psi_{occ}) = (N - 2)/2$  and  $z_2 = \text{Dim}(\bar{\Psi}_{occ}) = (N - 2)/2 + 2$ . There are also two 1/2-in-gap states for this case. If  $B(0) = -B(\pi)$ , we have  $z_1 = z_2 = N/2$ , which means no protected 1/2-in-gap states.

For more than one filled band, we repeat the process for each filled band, and find that  $z_1 = \text{Dim}(\Psi_{occ})$  is the total number of +1 in the diagonal of  $\mathcal{B}(0)$  and  $\mathcal{B}(\pi)$ , and  $z_2 = \text{Dim}(\bar{\Psi}_{occ})$  is the total number of -1 in the diagonal of  $\mathcal{B}(0)$  and  $\mathcal{B}(\pi)$ :

$$\begin{aligned}z_1 &= n_+(0) + n_+(\pi) = \sum_{k_{inv}} n_+(k_{inv}) \\ z_2 &= n_-(0) + n_-(\pi) = \sum_{k_{inv}} n_-(k_{inv}).\end{aligned}\tag{32}$$

Substituting these expressions into Eq.(26), we recover Eq.(12). (The sewing matrix used here is different from the one used in Ref.[32] due to different choice of inversion center.)

### C. Protection of in-gap states against inversion symmetric and weak disorder

We would like to understand how these new quantum numbers,  $z_1$  and  $z_2$  will change when we consider a disordered system. Since our previous discussion in Sec.II(A) did not rely on translational invariance the answer will not change so far as the disorders are inversion symmetric. We begin by assuming a form for the on-site potential which preserves the inversion symmetry  $\hat{V} = \sum_{\alpha r} V(r)c_{\alpha}^{\dagger}(r)c_{\alpha}(r)$  satisfying  $V(r) = V(N+1-r)$ , then we have  $\hat{P}\hat{V}\hat{P} = \hat{V}$ . Therefore,

$$\langle \bar{\psi}_m | \hat{V} | \psi_n \rangle = 0.\tag{33}$$

This entails that if  $\hat{V}$  is tuned adiabatically, an even parity state  $|\psi_m\rangle \in \Psi$  will always remain in the even parity subspace  $\Psi$ , so the total number of occupied states of even parity,  $z_1$ , does not change during the process. Similar argument gives that  $z_2$  also remains unchanged during this adiabatic process. The many-body ground state for  $\hat{H} + \hat{V}$  will be the same as the state evolved from the many-body ground state for  $\hat{H}$ , if there is no level crossing in the adiabatic process. Therefore, we

conclude that, as far as  $|V|$  is small compared with the bulk gap, the numbers of occupied states  $z_1$  and  $z_2$  in even and odd subspaces are the same as  $z_1$  and  $z_2$  in the homogeneous system. Thus, the number of 1/2-in-gap states,  $|z_1 - z_2|$ , is unchanged and the 1/2-in-gap states are robust against any inversion symmetric disorder, and do not depend on translational symmetry in general.

Before ending our revisit to the 1D inversion invariant insulators, we must emphasize the following two points that give us insights for extending our work on inversion invariant insulators to  $C_n$  invariant insulators. First, we want to choose a subsystem that is conjugal to its complementary subsystem with respect to the point group symmetry, because in this case, the correlation matrix decomposes into a linear average of projectors. Second, we can define an integer for each 1D representation of the point group, which is the total number of occupied states that transform according to that representation. The sum of these integers gives the total number of electrons, while their difference is related to the protected in-gap states in the entanglement spectrum.

## III. ENTANGLEMENT SPECTRUM OF INVERSION INVARIANT TOPOLOGICAL INSULATORS IN 2D AND 3D

The conclusion we arrived at for the 1D insulators with inversion symmetry can be easily generalized to higher dimensions. In these insulators, one can also define the  $Z^2$  index as the number of occupied states with even and odd parities:  $(z_1, z_2) = (\text{Dim}(\Psi_{occ}), \text{Dim}(\bar{\Psi}_{occ}))$ . The sum of its two components  $z_1 + z_2$  gives the total number of electrons and their difference,  $|z_1 - z_2|$ , is the number of 1/2-in-gap states in the entanglement spectrum for any inversion symmetric cut. Similar to the symmetric cut in 1D, one defines a symmetric cut in 2D (3D) as a proper subset of sites called  $A$  that has exactly one half of all sites, and if site  $\mathbf{r} \in A$ , then  $\mathbf{Pr} \notin A$ . The detailed proof of  $N_{1/2} = |z_1 - z_2|$  is omitted because it takes exactly the same steps, which we simply summarize here as: 1. Write the correlation matrix  $C$  as the linear average of two projector matrices  $D$  and  $\bar{D}$ . 2. Show that the number of 1/2-in-gap states is the difference between the number of unity eigenvalues of  $D$  and  $\bar{D}$ . 3. Show that  $\text{dim}(D) = \text{Dim}(\Psi_{occ}) \equiv z_1$  and  $\text{dim}(\bar{D}) = \text{Dim}(\bar{\Psi}_{occ}) \equiv z_2$ .

We can easily count  $\text{Dim}(\Psi_{occ})$  and  $\text{Dim}(\bar{\Psi}_{occ})$  in presence of translational invariance. For each pair of  $(\mathbf{k}, -\mathbf{k})$  where  $\mathbf{k} \neq \mathbf{k}_{inv}$ , there are for each band exactly one parity odd and one parity even eigenstate, thus contributing equally to  $\text{Dim}(\Psi_{occ})$  and  $\text{Dim}(\bar{\Psi}_{occ})$ . The contribution from all these pairs is  $(N - 4)N_{occ}/2$  in 2D and  $(N - 8)N_{occ}/2$  in 3D, as there are four and eight  $\mathbf{k}_{inv}$ 's in 2D and 3D respectively. In the following, we will neglect the contribution from all of the aforementioned pairs in the  $Z^2$  index, as they are completely determined by the total number of sites and the total number of bands (or particles). As in 1D, the interesting part is



contributed by eigenstates at  $\mathbf{k}_{inv}$ 's, because the eigenstate at  $\mathbf{k}_{inv}$  on each band is either parity even or parity odd. The contribution from these  $\mathbf{k}_{inv}$ 's to the  $Z^2$  index is given by

$$\begin{aligned} Dim(\Psi_{occ.}) &= \sum_{\mathbf{k}_{inv}} n^+(\mathbf{k}_{inv}), \\ Dim(\bar{\Psi}_{occ.}) &= \sum_{\mathbf{k}_{inv}} n^-(\mathbf{k}_{inv}). \end{aligned} \quad (34)$$

The  $Z^2$  index obtained in a translationally invariant system remains unchanged in the disordered system as long as the disorder potential preserve inversion symmetry, as mentioned previously. We may generally state that the number of 1/2-in-gap states in the entanglement spectrum for any inversion symmetric cut with or without weak disorder is given by the same Eq.(12), with the only difference that now  $\mathbf{k}_{inv}$  runs over four and eight points in 2D and 3D respectively.

In addition to giving the number of 1/2-in-gap states, the  $Z^2$  index also gives the parity of Chern number in 2D. In fact, we have

$$\begin{aligned} (-1)^{(z_1 - z_2)/2} &= (-1)^{\sum_{\mathbf{k}_{inv}} n^-(\mathbf{k}_{inv})} \\ &= (-1)^C. \end{aligned} \quad (35)$$

In the last equality, we have applied Eq.(92,93) in Ref.[32], which relate the Chern number to the inversion eigenvalues at all inversion invariant points in BZ.

In 2D and 3D, if in addition to inversion symmetry, we also have time-reversal symmetry, every band is at least doubly degenerate (for spinful fermions). And since  $[\hat{T}, \hat{P}] = 0$ , the two degenerate bands must have equal inversion eigenvalue at  $\mathbf{k}_{inv}$ 's. In this case  $(z_1 - z_2)/2$  must be an even number and one can further have

$$\begin{aligned} (-1)^{(z_1 - z_2)/4} &= (-1)^{\sum_{\mathbf{k}_{inv}} n^-(\mathbf{k}_{inv})/2} \\ &= (-1)^{\gamma_0}, \end{aligned} \quad (36)$$

where  $\gamma_0$  is the  $Z_2$  index of a TRI insulator.

As a final remark, the  $Z^2$ -index only concerns the total number of occupied states with even/odd parities and therefore does not exhaust the topological classifications. For example, an inversion symmetric insulator with only one filled band having  $\mathcal{B}(\Gamma) = \mathcal{B}(X) = -\mathcal{B}(Y) = -\mathcal{B}(M) = 1$  is topologically different from another one with  $\mathcal{B}(\Gamma) = \mathcal{B}(M) = -\mathcal{B}(X) = -\mathcal{B}(Y) = 1$ , though they have the same  $Z^2$ -indices.

#### IV. ENTANGLEMENT SPECTRUM IN $C_n$ INVARIANT INSULATORS

We are already aware that we can define a  $Z^2$  index for inversion invariant insulators. This index represents the number of occupied eigenstates in each of the two 1D representations of the inversion point group. In total, there are  $n$  1D representations of point group  $C_n$ , in each of which  $C_n$  is represented by a root of the

equation  $x^n = (-1)^F$ . We denote the  $m$ th root by  $x_m = e^{i(F+2(m-1))\pi/n}$ . This allows us to define a  $Z^n$  index  $(z_1, \dots, z_n)$ , in which  $z_m$  is the number of occupied eigenstates in a subspace  $\Psi^{(x_m)}$

$$z_m = Dim(\Psi_{occ}^{(x_m)}). \quad (37)$$

Any state  $|\psi\rangle$  in  $\Psi^{(x_m)}$  satisfies  $\hat{C}_n|\psi\rangle = x_m|\psi\rangle$ , or, in other words, transforms according to the  $m$ -th representation of group  $C_n$ .

Yet it is not clear in what manner this  $Z^n$  index is related to the number of protected in-gap states in the single particle spectrum. To answer this, we must define what a symmetric cut is in a  $C_n$  invariant insulator, analogous to the cut defined with inversion invariant insulators, because a given entanglement spectrum is associated with a specific real space cut. Due to the complexity of  $C_n$  compared with the inversion point group, there are more than one type of symmetric cuts. Two integers  $m_1$  and  $m_2$ , both being factors  $n$ , describe a symmetric cut denoted by  $A_{1/m_1}^{m_2}$  that satisfies (a) the number of sites in  $A$  is exactly  $N/m_1$ . (b) Subset  $A$  is invariant under  $m_2$ -fold rotation, or symbolically,  $C_{m_2}\mathbf{r} \in A, \forall \mathbf{r} \in A$  and (c) the subset  $A$ , together with all copies of  $A$  obtained from acting  $C_{m_1 m_2}$  rotation on  $A$  for  $k < m_1$  times, constitute the whole system, or symbolically,  $\cup_{k=0, \dots, m_1-1} C_{m_1 m_2}^k A = L$ , where  $L$  represents the whole lattice. Notice that properties (a) and (c) together imply (d)  $C_{m_1 m_2}^{k_1} A \cap C_{m_1 m_2}^{k_2} A = \emptyset$ , if  $k_1 \neq k_2 \bmod m_1$ , while  $C_{m_1 m_2}^{m_1} = C_{m_2}$ . In other words, an  $m_1 m_2$ -fold rotation on  $A$  generates other equivalent subsystems that are invariant under  $m_2$ -fold rotation. For  $n = 2, 3, 4, 6$ , all possible combinations of  $(m_1, m_2)$  are listed with typical examples of every type of cuts in the real space in Table I. Additionally, we note that for a fixed  $n$ , there is a one-to-one mapping between one type of symmetric cuts and one coset decomposition of a subgroup of  $C_n$ .

In the last column of Table I, the number of the protected in-gap states is given in terms of  $Z^n$  indices. This gives the *least* number of eigenvalues in the entanglement spectrum that are constrained in the range  $[1/m_1, 1 - 1/m_1]$ . If  $m_1 = 2$ , they are all at exactly 1/2 as we have seen in the inversion invariant insulators. But when  $m_1 > 2$ , we do not have protected states at exactly 1/2, but states that are prevented from moving outside the range of  $[1/m_1, 1 - 1/m_1]$ .

To prove these expressions, we first show that the entanglement matrix calculated with an  $A_{1/m_1}^{m_2}$ ,  $C(A_{1/m_1}^{m_2})$ , can be block-diagonalized into  $m_2$  blocks. Since  $A$  is invariant under  $C_{m_2}$ , we can define a proper set of  $A$  with exactly  $N_A/m_2$  sites,  $A_0$ , such that  $C_{m_2}^p \mathbf{r} \notin A_0$  if  $\mathbf{r} \in A_0$  for  $p = 1, \dots, m_2 - 1$ . For every site  $\mathbf{r} \in A_0$  we define the following  $m_2$  symmetry adapted operators

$$d_{y_q, \alpha}(\mathbf{r}) = \frac{1}{\sqrt{m_2}} \sum_{p=1, \dots, m_2} (y_q^*)^p \hat{C}_{m_2}^p c_\alpha(\mathbf{r}) \hat{C}_{m_2}^{-p}, \quad (38)$$

where  $y_q = \exp[i(2q + F - 2)\pi/m_2]$  is the  $q$ th root of the  $m_2$  roots of  $e^{iF\pi}$  ( $q = 1, \dots, m_2$ ). (For an intuitive schematic of this definition, see Fig.3.)

$$\begin{aligned}
A_{1/2}^3 &= \text{Diagram 1} & A_0 &= \text{Diagram 2} \\
d_{e^{iF\pi/3}} &= \frac{1}{\sqrt{3}} (\text{Diagram 3} + \text{Diagram 4} + \text{Diagram 5}) \\
d_{e^{i(F+2)\pi/3}} &= \frac{1}{\sqrt{3}} (\text{Diagram 6} + e^{-i\frac{2\pi}{3}} \text{Diagram 7} + e^{-i\frac{4\pi}{3}} \text{Diagram 8}) \\
d_{e^{i(F+4)\pi/3}} &= \frac{1}{\sqrt{3}} (\text{Diagram 9} + e^{-i\frac{4\pi}{3}} \text{Diagram 10} + e^{-i\frac{2\pi}{3}} \text{Diagram 11})
\end{aligned}$$

FIG. 3. A picturesque expression of the definition of operators  $d_{y_q, \alpha}$ . We use  $A_{1/2}^3$  as an example, and show how the operators in  $A$  can be decomposed into three sectors with different eigenvalues of  $C_3$ .

These new operators have two properties:  $\{d_{y_q, \alpha}(\mathbf{r}), d_{y_{q'}, \beta}^\dagger(\mathbf{r}')\} = \delta_{qq'} \delta_{\alpha\beta} \delta_{\mathbf{r}\mathbf{r}'}$  and

$$\hat{C}_{m_2} d_{y_q, \alpha}(\mathbf{r}) \hat{C}_{m_2}^{-1} = y_q d_{y_q, \alpha}(\mathbf{r}). \quad (39)$$

The first property ensures that the original operators  $c_\alpha(\mathbf{r} \in A)$  can be linked to  $d_{y_q, \alpha}(\mathbf{r} \in A_0)$  by a unitary transform, thus preserving all eigenvalues of the correlation matrix. The second property leads to

$$\langle d_{y_{q_1}, \alpha}^\dagger(\mathbf{r}) d_{y_{q_2}, \beta}(\mathbf{r}') \rangle \propto \delta_{q_1 q_2}. \quad (40)$$

To see this, we first notice that the ground state is  $C_{m_2}$ -invariant, i.e.,  $\hat{C}_{m_2}|0\rangle = e^{i\theta}|0\rangle$ , where  $\theta$  is some arbitrary angle. Then use Eq.(39):

$$\begin{aligned}
&\langle d_{y_{q_1}, \alpha}^\dagger(\mathbf{r}) d_{y_{q_2}, \beta}(\mathbf{r}') \rangle \\
&= \langle 0 | \hat{C}_{m_2}^{-1} (\hat{C}_{m_2} d_{y_{q_1}, \alpha}^\dagger(\mathbf{r}) \hat{C}_{m_2}^{-1}) (\hat{C}_{m_2} d_{y_{q_2}, \beta}(\mathbf{r}') \hat{C}_{m_2}^{-1}) \hat{C}_{m_2} | 0 \rangle \\
&= y_{q_1}^{-1} y_{q_2} \langle d_{y_{q_1}, \alpha}^\dagger(\mathbf{r}) d_{y_{q_2}, \beta}(\mathbf{r}') \rangle.
\end{aligned} \quad (41)$$

If  $q_1 \neq q_2$ , from the above equation, we have  $\langle d_{y_{q_1}, \alpha}^\dagger(\mathbf{r}) d_{y_{q_2}, \beta}(\mathbf{r}') \rangle = 0$ .

From these we know that using the basis of  $d_{y_q, \alpha}(\mathbf{r} \in A_0)$ , the correlation matrix is block diagonalized into  $m_2$  blocks, i.e.,

$$C(A) = C^{(1)}(A_0) \oplus C^{(2)}(A_0) \dots \oplus C^{(m_2)}(A_0), \quad (42)$$

where

$$C_{\mathbf{r}\alpha, \mathbf{r}'\beta}^{(p)}(A_0) = \langle d_{y_p, \alpha}^\dagger(\mathbf{r}) d_{y_p, \beta}(\mathbf{r}') \rangle. \quad (43)$$

Using this result, the study of in-gap states in  $C(A)$  reduces to the study of in-gap states in each block  $C^{(p)}(A_0)$ . When we have calculated the number of in-gap states,  $N_{mid}^{(q)}$ , for each block, we can simply add up all these numbers and obtain

$$N_{mid}(A_{1/m_1}^{m_2}) = \sum_{q=1, \dots, m_2} N_{mid}^{(p)} \quad (44)$$

for the total number of in-gap states.

From here we only focus on one block,  $C^{(p)}(A_0)$ , of the block diagonalized correlation matrix. In Appendix C, we prove that  $C^{(p)}(A_0)$  can be written as a linear average of  $m_1$  projectors

$$C^{(p)}(A_0) = \frac{1}{m_1} (D_1^{(p)} + \dots + D_{m_1}^{(p)}), \quad (45)$$

where  $D_r^{(p)}$  is a projector, projecting any state into a subspace  $\Phi_r^{(p)}$ , in which every state satisfies these two conditions: (i) it is an eigenvector of  $\hat{C}_{m_2}$  of eigenvalue  $y_p$ , and (ii) it is an eigenvector of  $\hat{C}_{m_1 m_2}$  with eigenvalue  $\lambda_r^{(p)}$ . Notice that since  $\hat{C}_{m_2} = \hat{C}_{m_1 m_2}^{m_1}$ , we have the constraint  $(\lambda_r^{(p)})^{m_1} = y_p$ . Solve it and we have  $\lambda_r^{(p)} = \exp[i\pi(F + 2((r-1)m_2 + p - 1))/(m_1 m_2)]$  for  $r = 1, \dots, m_1$ . Given an eigenstate of  $C_n$  of eigenvalue  $x_{m=1, \dots, n}$ , if it is in  $\Phi_r^{(p)}$ , using  $\hat{C}_{m_1 m_2} = \hat{C}_n^{\frac{n}{m_1 m_2}}$   $x_m$  must satisfy  $x_m^{\frac{n}{m_1 m_2}} = \lambda_r^{(p)}$ .

In Appendix C, we prove that the number of non-zero eigenvalues of  $D_r^{(p)}$  is

$$\begin{aligned}
\dim(D_r^{(p)}) &= \dim(\cup_{x_m} \{\Psi_{occ}^{(x_m)} | x_m^{n/(m_1 m_2)} = e^{i2\pi((r-1)m_2 + p - 1 + F/2)/(m_1 m_2)}\}) \\
&= \sum_{x_m^{n/(m_1 m_2)} = e^{i2\pi((r-1)m_2 + p - 1 + F/2)/(m_1 m_2)}} z_m.
\end{aligned} \quad (46)$$

Recall that in inversion invariant insulators, the difference between  $\dim(D)$  and  $\dim(\bar{D})$  corresponds to the number of in-gap states; here we have a very similar relation. The block of the correlation matrix  $C^{(p)}$  has

$$N_{mid}^{(p)}(A_0) = \max_{i,j=1, \dots, m_1} |\dim(D_i^{(p)}) - \dim(D_j^{(p)})| \quad (47)$$

eigenstates within the range  $[1/m_1, 1 - 1/m_1]$  (see Appendix D for proof). Therefore the total number of in-gap states is the sum of the number of in-gap states in

each block:

$$N_{mid}(A_{1/m_1}^{m_2}) = \sum_{p=1, \dots, m_2} \max_{i,j=1, \dots, m_1} |\dim(D_i^{(p)}) - \dim(D_j^{(p)})|. \quad (48)$$

Substituting Eq.(46), which relates  $\dim(D_r^{(p)})$  to  $Z^n$ -index, into Eq.(48), we obtain the expression of the total number of protected in-gap states for the symmetric cut

$A_{1/m_1}^{m_2}$  as shown in the last column of Table I.

The discussion above applies to systems with and without translational invariance. In the case when translational invariance is present, we desire to elucidate how the  $Z^n$  index related to the eigenvalues of symmetry operators at high symmetry points. To understand the nature of this link, we calculate the  $Z^n$  index given a single particle Hamiltonian in  $\mathbf{k}$ -space, or in the homogeneous limit with periodic boundaries. To be exact, we want to count the dimension of the linear subspace  $\Psi_{occ}^{(x_m)}$ , where  $x_q = \exp(i(F + 2q - 2)\pi/n)$  for  $q = 1, \dots, n$ . First we write symmetry adapted operators in terms of annihilation operators the Bloch states,  $\gamma_m(\mathbf{k})$ :

---


$$\psi_{x_q, m}(\mathbf{k}) = \frac{1}{\sqrt{n}} \sum_{p=0, \dots, n-1} (x_q^*)^p \hat{C}_n^p \gamma_m(\mathbf{k}) \hat{C}_n^{-p}, \quad (49)$$


---

$$\psi_{x_q, m}(\mathbf{k}_0) = \frac{1}{\sqrt{n}} \sum_{t=0, \dots, n'-1} (x_q^{*n/n'}) (\mathcal{B}_{n'}(\mathbf{k}_0))_{mm}^t \sum_{p=0, \dots, n/n'-1} (x_q^*)^p \hat{C}_n^p \gamma_m(\mathbf{k}_0) \hat{C}_n^{-p}, \quad (50)$$

where  $\mathcal{B}_{n'}$  is the sewing matrix corresponding to the  $n'$ -fold rotation. This expression is non-zero if and only if  $(\mathcal{B}_{n'}(\mathbf{k}_0))_{mm} = x_q^{n/n'}$ . It can be physically understood as follows: since  $\hat{C}_{n'} = \hat{C}_n^{n/n'}$ , and  $(\mathcal{B}_{n'}(\mathbf{k}_0))_{mm}$  is the eigenvalue of  $\hat{C}_{n'}$  for the state at  $\mathbf{k}_0$ , it must be the  $n/n'$ -th power of  $x_q$  to be in the subspace of  $\Psi^{(x_q)}$ . Furthermore, in the BZ there are exactly  $n/n'$  points that are  $C_{n'}$  symmetric but not  $C_n$  symmetric ( $n' < n$ ), and therefore there are  $n/n'$  orthogonal states, the linear combinations of which give the  $n/n'$  eigenstates of  $\hat{C}_n$  with eigenvalues satisfying  $x_q^{n/n'} = (\mathcal{B}_{n'}(\mathbf{k}_0))_{mm}$ .

It is straightforward to show  $\hat{C}_n \psi_{x_q, m}(\mathbf{k}) \hat{C}_n^{-1} = x_q \psi_{\epsilon, m}(\mathbf{k})$ . Therefore if  $\psi_{x_q, m}(\mathbf{k}) \neq 0$  and  $m \in occ$ , then state  $\psi_{x_q, m}^\dagger(\mathbf{k})|0\rangle$  belongs to subspace  $\Psi_{occ}^{(x_q)}$ , contributing +1 to  $z_q$ . With Eq.(50), the counting is therefore completely determined by the sewing matrices at these high symmetry points, or the number of each eigenvalue of  $\hat{C}_{n'}$  at these points. The result for  $n = 2, 3, 4, 6$  are listed in Table II. In Appendix F, we give an example of

---

where  $\gamma_m(\mathbf{k})$  is the annihilation operator of the Bloch state at  $\mathbf{k}$  on the  $m$ th band. This definition is a little different from that for  $d_{y_q, \alpha}$  in Eq.(38), though they transform in similar ways under rotation: (i)  $d_{y_q, \alpha}$  are to block diagonalize the correlation matrix, while  $\psi_{x_q, m}$ 's in general cannot do this; (ii)  $d_{y_q, \alpha}(\mathbf{r})$ 's are *not* annihilation operators of single particle eigenstates, while  $\psi_{x_q, m}(\mathbf{k})$ 's are generically annihilation operators of a single particle eigenstate; (iii) as we will see later,  $\psi_{x_q, m}$  can be *zero* at points of symmetry while  $d_{y_q, \alpha}$  is always nonzero. We notice that for a generic  $\mathbf{k}$  that is not invariant under any subgroup of  $C_n$ ,  $\{\psi_{x_{q_1}, m}(\mathbf{k}), \psi_{x_{q_2}, m'}^\dagger(\mathbf{k})\} = \delta_{q_1 q_2} \delta_{mm'}$ . Any one of these  $\mathbf{k}$  points on each band contributes +1 to every component of the  $Z^n$  index. Again this contribution is ignored as it is completely determined by the total number of particles and total number of bands. But if  $\mathbf{k}_0$  is invariant under a subgroup  $C_{n'}$  of  $C_n$ , i.e.,  $C_{n'} \mathbf{k}_0 = \mathbf{k}_0$ , we have, applying Eq.(B5) to Eq.(49),

---

deriving the row with  $n = 6, m_1 = 2, m_2 = 3$  in Table II, following the general rules given above.

Substituting the dimension counting into Eq.(48), we have finally linked the number of in-gap states in a  $C_n$  invariant insulator with a symmetric cut  $A_{1/m_1}^{m_2}$  to the number of different eigenvalues of  $\hat{C}_m$  ( $m$  dividing  $n$ ) at all high symmetry points. Finally, it should be noted that although we use the sewing matrix, defined only for a homogeneous system, to calculate the  $Z^n$  index, the components of the index do not change under any adiabatic shift of Hamiltonian including addition of  $C_n$  symmetric disorder potential. Therefore, as far as the disorder strength is small compared with the insulating gap, the number of in-gap states given in Table II is protected from changing only by  $C_n$  symmetry but not by translational symmetry, and is uniquely determined by the  $Z^n$  index.

In Fig.4, we give two specific examples of a  $C_4$  invariant 2D insulator showing protected 1/2-in-gap states in the entanglement spectrum, both with symmetric cuts of the type  $A_{1/2}^2$ . The system is described by a tight-binding model with two bands:

$$H(k_x, k_y) = (m - \cos(k_x) - \cos(k_y))\sigma_z + \sin(k_x)\sigma_x + \sin(k_y)\sigma_y, \quad (51)$$


---

where  $m$  is a real tunable parameter. This Hamiltonian is  $C_4$ -invariant,  $C_4(\mathbf{k})H(k_x, k_y)C_4^{-1}(\mathbf{k}) = H(-k_y, k_x)$ , with  $C_4(\mathbf{k}) = \exp(i\sigma_z \pi/4) e^{ik_x}$ . Mark that  $C_4$ -matrix

is  $\mathbf{k}$ -dependent since the rotation center is not chosen at a lattice site, but the center of a plaquette. Using  $C_2(k_x, k_y) = C_4(-k_y, k_x)C_4(k_x, k_y)$ , we have  $C_2(\mathbf{k}) =$

$n$	$m_1$	$m_2$	Typical examples	Minimal number of midgap states in the entanglement spectrum
2	2	1		$ z_1 - z_2 $
3	3	1		$\max_{i,j=1,2,3}  z_i - z_j $
4	2	1		$ z_1 + z_3 - z_2 - z_4 $
	2	2		$ z_1 - z_3  +  z_2 - z_4 $
	4	1		$\max_{i,j=1,2,3,4}  z_i - z_j $
6	2	1		$ z_1 + z_3 + z_5 - z_2 - z_4 - z_6 $
	2	3		$ z_1 - z_4  +  z_2 - z_5  +  z_3 - z_6 $
	3	1		$\max( z_1 + z_4 - z_2 - z_5 ,  z_1 + z_4 - z_3 - z_6 ,  z_2 + z_5 - z_3 - z_6 )$
	3	2		$\max( z_1 - z_3 ,  z_1 - z_5 ,  z_3 - z_5 ) + \max( z_2 - z_4 ,  z_2 - z_6 ,  z_4 - z_6 )$
	6	1		$\max_{i,j=1,\dots,6}  z_i - z_j $

TABLE I. All possible symmetric cuts  $A_{1/m_1}^{m_2}$  in  $C_n$  invariant insulators. In the fourth column, examples of each case are shown and in the last column, the minimal numbers of in-gap states in the entanglement spectrum are represented in terms of the  $Z^n$ -index.

$i\sigma_z e^{i(k_x - k_y)}$ . We assume that the system is half-filled. There are two  $C_4$ -symmetric points  $\Gamma$  and  $M$ , where the Hamiltonians are  $H(\Gamma) = (m - 2)\sigma_z$  and  $H(M) = (m + 2)\sigma_z$ , respectively. If  $m > 2$ , the occupied state is  $(0, 1)^T$  at both points, giving  $\mathcal{B}(\Gamma) = e^{-i\pi/4}$  and  $\mathcal{B}(M) = -e^{-i\pi/4}$ . At  $X$  or  $Y$ ,  $H(X/Y) = m\sigma_z$ , and the occupied state is still  $(0, 1)^T$ , giving  $\mathcal{B}(X/Y) = i$ . According to Table II, we find  $z_1 = z_2 = z_3 = z_4 = 1$ . For  $m < -2$ , similarly, we can similarly derive  $z_{1,2,3,4} = 1$ . For  $0 < m < 2$ , at  $\Gamma$ , the occupied state becomes  $(1, 0)^T$ , contributing  $\mathcal{B}(\Gamma) = e^{i\pi/4}$ , while for  $M$  and  $X$ , the occupied state is still  $(0, 1)^T$ . In this case we have  $z_1 = 2$ ,  $z_2 = z_3 = 1$  and  $z_4 = 0$ . Finally for  $-2 < m < 0$ , one can similarly derive  $z_1 = z_4 = 1$ ,  $z_2 = 2$  and  $z_3 = 0$ . Summarizing all possibilities, we have  $z_1 = 1 - \text{sgn}(m - 2)/2 + \text{sgn}(m)$ ,  $z_2 = 1 + \text{sgn}(m + 2)/2 - \text{sgn}(m)/2$ ,  $z_3 = 1 - \text{sgn}(m + 2) + \text{sgn}(m)/2$  and  $z_4 = 1 + \text{sgn}(m - 2)/2 - \text{sgn}(m)$ . From Table I, we know that with this type of symmetric cut, the number of  $1/2$  in-gap states are  $N_{1/2} = |z_1 - z_3| + |z_2 - z_4| = 2|\text{sgn}(m + 2) - \text{sgn}(m - 2)|$ . If  $|m| > 2$ , then there is no protected in-gap state and if  $|m| < 2$ , there are always two degenerate entanglement eigenvalues at exactly  $1/2$ . In Fig.4, we confirm this result by calculating the entanglement spectrum with or without  $C_4$  invariant impurities. In the top two figures of Fig.4, we can see that although the details of the spectrum depends on the specific choice of the real space cut, the presence of the

doubly degenerate protected in-gap states is common to both cuts. In the bottom two figures of Fig.4, we can see that random impurities in general push the eigenvalues to the two ends of the spectrum as the electronic states become more localized, the doubly degenerate in-gap states always remain at  $1/2$ .

## V. DISCUSSION

### A. Entanglement spectrum of 3D insulators with PGS

The major focus of the paper is 2D insulators with  $n$ -fold rotation symmetry; however, the results obtained can be promptly extended to 3D insulators with  $n$ -fold rotation symmetry. This is because a 3D  $C_n$  insulator can be considered as a 2D  $C_n$  insulator parameterized by a continuous  $k_z$  along the rotation axis. For every  $k_z$  we can find its number of in-gap states in the entanglement spectrum and, as we require the 3D system to be insulating, this number of in-gap states must be a constant for different  $k_z$ . Therefore, studying the system at one  $k_z$ -slice as a 2D insulator is sufficient to understanding the symmetry protected in-gap states in the entanglement spectrum in the 3D  $C_n$ -invariant system.

For 3D insulators with other than  $C_n$  PGS, their entanglement spectra in principle show other symmetry pro-



$n = 2$	$Dim(\Psi_{occ}^{(e^{iF\pi/2})})$	$n_2^{(e^{iF\pi/2})}(\Gamma) + n_2^{(e^{iF\pi/2})}(X) + n_2^{(e^{iF\pi/2})}(Y) + n_2^{(e^{iF\pi/2})}(M)$
	$Dim(\Psi_{occ}^{(e^{i(F+2)\pi/2})})$	$n_2^{(e^{i(F+2)\pi/2})}(\Gamma) + n_2^{(e^{i(F+2)\pi/2})}(X) + n_2^{(e^{i(F+2)\pi/2})}(Y) + n_2^{(e^{i(F+2)\pi/2})}(M)$
$n = 3$	$Dim(\Psi_{occ}^{(e^{iF\pi/3})})$	$n_3^{(e^{iF\pi/3})}(\Gamma) + n_3^{(e^{iF\pi/3})}(K) + n_3^{(e^{iF\pi/3})}(K')$
	$Dim(\Psi_{occ}^{(e^{i(F+2)\pi/3})})$	$n_3^{(e^{iF\pi/3})}(\Gamma) + n_3^{(e^{i(F+2)\pi/3})}(K) + n_3^{(e^{i(F+2)\pi/3})}(K')$
	$Dim(\Psi_{occ}^{(e^{i5F\pi/3})})$	$n_3^{(e^{i(F+4)\pi/3})}(\Gamma) + n_3^{(e^{i(F+4)\pi/3})}(K) + n_3^{(e^{i(F+4)\pi/3})}(K')$
$n = 4$	$Dim(\Psi_{occ}^{(e^{iF\pi/4})})$	$n_4^{(e^{iF\pi/4})}(\Gamma) + n_4^{(e^{iF\pi/4})}(M) + n_2^{(e^{iF\pi/2})}(X)$
	$Dim(\Psi_{occ}^{(e^{i(F+2)\pi/4})})$	$n_4^{(e^{i(F+2)\pi/4})}(\Gamma) + n_4^{(e^{i(F+2)\pi/4})}(M) + n_2^{(e^{i(F+2)\pi/2})}(X)$
	$Dim(\Psi_{occ}^{(e^{i(F+4)\pi/4})})$	$n_4^{(e^{i(F+4)\pi/4})}(\Gamma) + n_4^{(e^{i(F+4)\pi/4})}(M) + n_2^{(e^{iF\pi/2})}(X)$
	$Dim(\Psi_{occ}^{(e^{i(F+6)\pi/4})})$	$n_4^{(e^{i(F+6)\pi/4})}(\Gamma) + n_4^{(e^{i(F+6)\pi/4})}(M) + n_2^{(e^{i(F+2)\pi/2})}(X)$
	$Dim(\Psi_{occ}^{(e^{i5F\pi/4})})$	$n_4^{(e^{i(F+4)\pi/4})}(\Gamma) + n_4^{(e^{i(F+4)\pi/4})}(M) + n_2^{(e^{iF\pi/2})}(X)$
$n = 6$	$Dim(\Psi_{occ}^{(e^{iF\pi/6})})$	$n_6^{(e^{iF\pi/6})}(\Gamma) + n_3^{(e^{iF\pi/3})}(K) + n_2^{(e^{iF\pi/2})}(M)$
	$Dim(\Psi_{occ}^{(e^{i(F+2)\pi/6})})$	$n_6^{(e^{i(F+2)\pi/6})}(\Gamma) + n_3^{(e^{i(F+2)\pi/3})}(K) + n_2^{(e^{i(F+2)\pi/2})}(M)$
	$Dim(\Psi_{occ}^{(e^{i(F+4)\pi/6})})$	$n_6^{(e^{i(F+4)\pi/6})}(\Gamma) + n_3^{(e^{i(F+4)\pi/3})}(K) + n_2^{(e^{iF\pi/2})}(M)$
	$Dim(\Psi_{occ}^{(e^{i(F+6)\pi/6})})$	$n_6^{(e^{i(F+6)\pi/6})}(\Gamma) + n_3^{(e^{iF\pi/3})}(K) + n_2^{(e^{i(F+2)\pi/2})}(M)$
	$Dim(\Psi_{occ}^{(e^{i(F+8)\pi/6})})$	$n_6^{(e^{i(F+8)\pi/6})}(\Gamma) + n_3^{(e^{i(F+2)\pi/3})}(K) + n_2^{(e^{iF\pi/2})}(M)$
	$Dim(\Psi_{occ}^{(e^{i(F+10)\pi/6})})$	$n_6^{(e^{i(F+10)\pi/6})}(\Gamma) + n_3^{(e^{i(F+4)\pi/3})}(K) + n_2^{(e^{i(F+2)\pi/2})}(M)$

TABLE II.  $Z^n$ -index in a  $C_n$  invariant insulator with translational symmetry. These indices are related to the number of all eigenvalues of the corresponding sewing matrices associated with subgroups of  $C_n$  at high symmetry points within the BZ.  $n_m^\epsilon$  represents the number of states with eigenvalue  $\epsilon$  of  $m$ -fold rotation  $C_m$ , where  $C_m$  is a subgroup of  $C_n$ .

tected features, and the number of in-gap states depend on other factors. Nevertheless, if the underlying PGS has only 1D representations, as  $C_n$  group does, one can still define the  $Z^{g_0}$ -index, where  $g_0$  is the order of the group, similar to the  $Z^n$ -index for  $C_n$ -invariant insulators. Then one can make exact statements about the number of in-gap states for properly defined symmetric real space cuts. If the group has higher-dimensional irreducible representations, it is yet to be studied how the occupation numbers of states belonging to these representation are related to the entanglement spectrum.

### B. Effects of weak interaction

Throughout the paper we have been studying systems without interactions. Now we discuss how weak interaction changes our results. By ‘weak’, we mean that the ground state of the interacting system can be obtained by adiabatically transforming the Slater determinant state or the ground state of the non-interacting system. This implies that the interacting ground state is both gapped and non-degenerate. Therefore, we are not considering any symmetry breaking state as such a state must be degenerate with at least another state related by the broken symmetry. Additionally, we are not considering fractional quantum Hall systems the ground states of which are degenerate. In this work, we assume the interacting ground state to be an insulator with a non-degenerate ground state and a many-body gap.

In an interacting insulator, single particle states lose their meaning, but the many-body ground state is still a 1D representation of  $C_n$ , giving a  $Z_n$  index. The relation between this  $Z_n$  index and the  $Z^n$  indices can be found by

noticing that the  $C_n$  rotation operator can be expressed as

$$\hat{C}_n = \exp[i\frac{2\pi}{n} \sum_{m=1, \dots, n, \alpha, \mathbf{r} \in L/n} (m-1) d_{x_m, \alpha}^\dagger(\mathbf{r}) d_{x_m, \alpha}(\mathbf{r})] \quad (52)$$

where

$$d_{x_m, \alpha}(r) = \frac{1}{\sqrt{n}} \sum_{p=1, \dots, n} x_m^p \hat{C}_n^p c_\alpha(\mathbf{r}) \hat{C}_n^{-p}, \quad (53)$$

and that

$$z_m = \langle \sum_{\alpha, \mathbf{r} \in L/n} d_{x_m, \alpha}^\dagger(\mathbf{r}) d_{x_m, \alpha}(\mathbf{r}) \rangle. \quad (54)$$

Since the ground state is an eigenstate of  $\hat{C}_n$ , we have

$$\begin{aligned} \langle \hat{C}_n \rangle &= \exp(i\frac{2\pi}{n} z) \\ &= \exp(i\frac{2\pi}{n} \sum_m (m-1) z_m), \end{aligned} \quad (55)$$

or

$$z = \sum_{m=1, \dots, n} (m-1) z_m \bmod n. \quad (56)$$

Eq.(56) shows how the  $Z^n$  classification in a non-interacting system downgrades to the  $Z_n$  classification in an interacting one. This indicates that a non-interacting insulator with a non-trivial  $Z^n$  index may have a trivial  $Z_n$  index in presence of weak interaction, and hence two topologically distinct non-interacting insulators may become indistinct as interaction is turned on, similar to the situation in superconductors<sup>47–49</sup>.

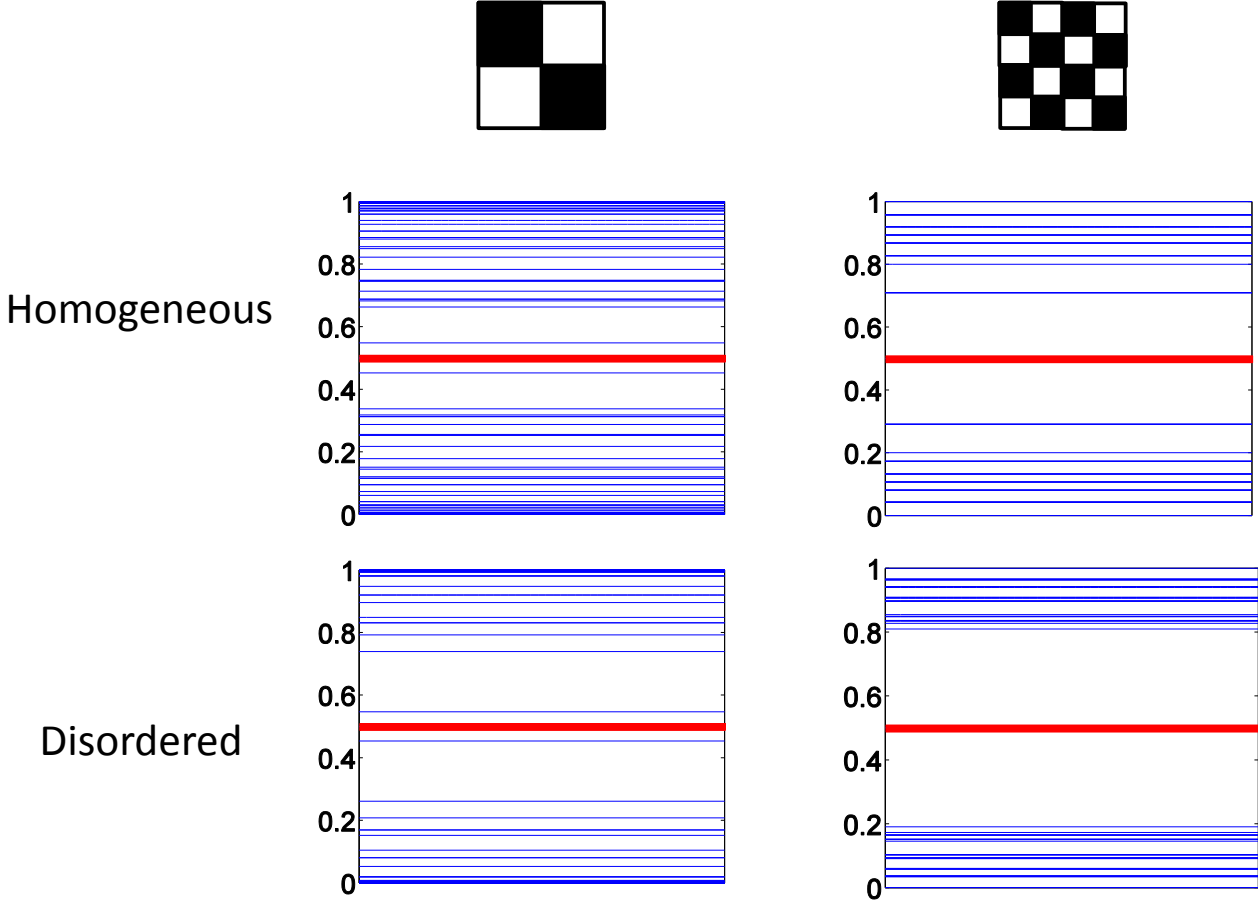


FIG. 4. Entanglement spectra in a  $C_4$  invariant insulator described by the Hamiltonian in Eq.(51) with  $m = 1.5$  are plotted with two different cuts of the type  $A_{1/2}^2$  on an 8-by-8 square lattice. The vertical axis is the entanglement eigenvalue and each level in the spectrum is represented by a horizontal line and Different columns are for different specific cuts; the first row shows the result without impurity (with translational symmetry) and the second with random disorders of strength  $\sim 0.4$  that are  $C_4$  symmetric. Red lines mark the position of eigenvalues at a value of exact  $1/2$ . The number of protected in-gap states is two in all cases.

Beyond the basic discussion presented here, the effects of interactions in point group symmetric topological insulators have many unanswered questions. A particularly important question is how the effects of interactions present in the changes in the entanglement spectrum. It is not clear how the in-gap states in the entanglement spectrum evolve as one turns on the interaction. We leave this question and other questions as to the role of interactions to future work.

## VI. CONCLUSION

In our analysis of the resultant entanglement spectrum in point group symmetric topological insulators, we have proven the relation between the  $1/2$ -in-gap states in the

entanglement spectrum in an inversion invariant insulator and the inversion eigenvalues at inversion invariant points. This was accomplished by noticing that, given any inversion invariant non-interacting insulator, one can always find a complete set of single particle eigenstates, of both the single particle Hamiltonian and inversion operator, and the numbers of occupied eigenstates that are odd/even under inversion are good quantum numbers. This results in a  $Z^2$ -index of the insulator. The sum of these two integers gives the total number of particles and their difference is exactly the number of  $1/2$ -in-gap states in the entanglement spectrum. We have extended this result from inversion invariant insulators to 2D invariant insulators with  $C_n$  invariance: The number of occupied states that transform according to each 1D representation of  $C_n$  is a good quantum number, resulting

in a  $Z^n$ -index. In a translationally invariant system, this  $Z^n$ -index can be calculated by counting the number of each  $C_m$  ( $m$  dividing  $n$ ) eigenvalue at high symmetry points. In the entanglement spectrum, the number of protected in-gap states is shown completely determined by the  $Z^n$  index with or without translational invariance in any  $C_n$  invariant insulator. Specially, if and only if all components in the  $Z^n$  index are equal, then the resultant number of protected in-gap states is zero. Finally, we briefly discussed how a similar analysis can be extended to 3D insulators with point groups having only 1D representations, and demonstrated that under weak interaction, the  $Z^n$ -index for a  $C_n$  invariant insulator re-

duces to a  $Z_n$ -index.

## ACKNOWLEDGMENTS

CF thanks A. Alexandradinata for helpful discussions. MJG acknowledges support from the AFOSR under grant FA9550-10-1-0459 and the ONR under grant N0014-11-1-0728 and a gift from the Intel Corporation. BAB was supported by NSF CAREER DMR-095242, ONR - N0014-11-1-0635, Darpa - N66001-11-1-4110 and David and Lucile Packard Foundation. MJG and BAB thank the Chinese Academy of Sciences for generous hosting.

## Appendix A: Subtleties in the definition of inversion operator in $\mathbf{k}$ -space

The relation between the number of in-gap states and the parities of occupied states at  $k = 0, \pi$  in a 1D inversion invariant insulator, Eq.(12), is completely consistent with Eq.(12) of Ref.32, although two expressions take different forms.

The difference lies in the definition of the inversion operator. While Ref.32 uses a  $k$ -independent  $\mathcal{P}$  as the matrix representation of inversion operator in  $k$ -space, in this paper we use a  $k$ -dependent  $\hat{\mathcal{P}}(k) = \mathcal{P}e^{ik}$  as in Ref.[34]. The factor  $e^{ik}$  comes from the fact that the inversion center is mid-bond, instead of on-site. Therefore, at  $k = \pi$ , the even/odd parity states in our paper correspond to odd/even states in Ref.32.

In 2D, the inversion operator in  $\mathbf{k}$ -space is  $\hat{\mathcal{P}}(\mathbf{k}) = \mathcal{P}e^{ik_x + ik_y}$  on a rectangular lattice, because the inversion center is at  $(\mathbf{a}_1 + \mathbf{a}_2)/2$ ; in 3D, we use  $\hat{\mathcal{P}}(\mathbf{k}) = \mathcal{P}e^{ik_x + ik_y + ik_z}$  as the inversion matrix in  $\mathbf{k}$ -space, because the inversion center is set at  $(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3)/2$ . An advantage of using such definitions is that the formula for the number of in-gap states in all dimensions takes exactly the same form:

$$N_{1/2} = \left| \sum_{\mathbf{k}_{inv}} n_+(\mathbf{k}_{inv}) - n_-(\mathbf{k}_{inv}) \right|. \quad (\text{A1})$$

## Appendix B: A basic property of the sewing matrix

In this Appendix we derive a simple property of a general sewing matrix associated with a PGS operation  $R$  we have applied to prove Eq.(29). The sewing matrix  $\mathcal{B}_R(\mathbf{k})$  is defined as

$$\mathcal{B}_{mn}(\mathbf{k}) = \langle u_m(R\mathbf{k}) | \hat{R} | u_n(\mathbf{k}) \rangle, \quad (\text{B1})$$

where  $m, n \in occ$ . Multiply both sides by  $|u_m(R\mathbf{k})\rangle$  then sum over  $m \in occ$ , we have

$$\mathcal{B}_{mn} |u_m(R\mathbf{k})\rangle = |u_m(R\mathbf{k})\rangle \langle u_m(R\mathbf{k}) | \hat{R} | u_n(\mathbf{k}) \rangle. \quad (\text{B2})$$

Since  $\hat{R}$  is a symmetry of the Hamiltonian, if one  $\hat{R} |u_n(\mathbf{k})\rangle$  must also be an occupied state at  $R\mathbf{k}$  if  $n \in occ$ , therefore  $\sum_{m \in occ} |u_m(R\mathbf{k})\rangle \langle u_m(R\mathbf{k})|$  can be replaced by the identity in the occupied bands. Eq.(B2) becomes

$$\hat{R} |u_n(\mathbf{k})\rangle = \mathcal{B}_{nn} |u_n(R\mathbf{k})\rangle. \quad (\text{B3})$$

Then we substitute

$$\begin{aligned} |u_n(\mathbf{k})\rangle &= \gamma_n^\dagger(\mathbf{k}) |0\rangle, \\ |u_m(R\mathbf{k})\rangle &= \gamma_m^\dagger(R\mathbf{k}) |0\rangle \end{aligned} \quad (\text{B4})$$

into Eq.(B3), and obtain

$$\hat{R} \gamma_n(\mathbf{k}) \hat{R}^{-1} = \mathcal{B}_{nn} \gamma_n(R\mathbf{k}). \quad (\text{B5})$$

### Appendix C: Decomposition of a correlation matrix into projectors

In Sec. IV, we have shown that the correlation matrix of a  $A_{1/m_1}^{m_2}$  cut can be block diagonalized into  $m_2$  blocks, each denoted by  $C^{(s)}(A)$ . In this appendix, we prove that this matrix can be further decomposed into the average of  $m_1$  projectors:

$$C^{(s)}(A) = \frac{1}{m_1} (D_1^{(s)} + \dots + D_{m_1}^{(s)}). \quad (C1)$$

We further prove that the number of unity eigenvalues in the projector  $D_i^{(s)}$  equals the number of occupied states the  $C_n$  eigenvalues  $\lambda$ 's of which satisfy  $\lambda^{n/(m_1 m_2)} = e^{i2\pi((i-1)m_2 + s - 1 + F/2)/(m_1 m_2)}$ .

Let us start from recalling that  $C_{i\alpha, j\beta}^{(s)} = \langle d_{y_s, \alpha}^\dagger(\mathbf{r}_i) d_{y_s, \beta}(\mathbf{r}_j) \rangle$ , where  $\mathbf{r}_{i,j} \in A_0$ . For each  $d_{y_s, \beta}(\mathbf{r}_i)$ , we further define  $m_1$  new operators

$$f_{\epsilon, \alpha}^{(s)}(\mathbf{r}_i) = \frac{1}{\sqrt{m_1}} \sum_{p=0, \dots, m_1-1} (\epsilon^*)^p C_{m_1 m_2}^p d_{\alpha}^{(s)}(\mathbf{r}_i) C_{m_1 m_2}^{\dagger p}, \quad (C2)$$

where  $\epsilon^{m_1} = \exp[i2\pi(\frac{s-1+F/2}{m_2})]$ . There are in total  $m_1$   $\epsilon$ 's that satisfy this condition, which are denoted by  $\epsilon_k = e^{i2\pi(km_2 + s + F/2)/(m_1 m_2)}$  for  $k = 1, \dots, m_1$ , and are all eigenvalues of operator  $\hat{C}_{m_1 m_2}$ . This is very similar to what we have done in the 1D inversion symmetric insulators (see Sec.II(A)): over there  $d_{\alpha}(\mathbf{r})$  and  $\bar{d}_{\alpha}(\mathbf{r})$  are combinations of  $c_{\alpha}(\mathbf{r})$  that are even and odd under inversion; here  $f_{\epsilon, \alpha}^{(s)}(\mathbf{r}_i)$  are linear combinations of  $d_{y_s, \alpha}(\mathbf{r})$  with different eigenvalues of  $C_{m_1 m_2}$ . These operators have two properties (a)  $C_{m_1 m_2} f_{\epsilon, \alpha}^{(s)}(\mathbf{r}_i) C_{m_1 m_2}^\dagger = \epsilon f_{\epsilon, \alpha}^{(s)}(\mathbf{r}_i)$  and (b)  $\{f_{\epsilon, \alpha}^{(s)}(\mathbf{r}_i), f_{\epsilon', \beta}^{(s)\dagger}(\mathbf{r}_j)\} = \delta_{\epsilon\epsilon'} \delta_{\alpha\beta} \delta_{\mathbf{r}_i \mathbf{r}_j}$ . From the first property, one derives  $\langle f_{\epsilon, \alpha}^{(s)\dagger}(\mathbf{r}_i) f_{\epsilon', \beta}^{(s)}(\mathbf{r}_j) \rangle = 0$  if  $\epsilon \neq \epsilon'$ . The derivation is similar to that of Eq.(40), only that here we use the fact that the ground state is an eigenstate of  $\hat{C}_{m_1 m_2}$ . Therefore the correlation matrix  $C^{(s)}(A_0)$  decomposes in the form of Eq.(45) where

$$(D_k^{(s)})_{i\alpha, j\beta} = \langle f_{\epsilon_k, \alpha}^{(s)\dagger}(\mathbf{r}_i) f_{\epsilon_k, \beta}^{(s)}(\mathbf{r}_j) \rangle. \quad (C3)$$

For simplicity, the superfix  $(s)$  will be suppressed and always implied throughout the remainder of this Appendix.

Now we need to prove that for every  $k = 1, \dots, m_1$ , matrix  $D_k$  is a projector, i.e.,  $(D_k)^2 = D_k$ . First we notice that  $C_{m_1 m_2}$  is a subgroup of  $C_n$  or  $C_n$  itself, so each single particle eigenstate has an eigenvalue of  $C_{m_1 m_2}$ . There are  $m_1 m_2$  eigenvalues of  $\hat{C}_{m_1 m_2}$ , dividing the single particle Hilbert space into  $m_1 m_2$  sectors, each of which is denoted by  $\Psi^\epsilon = \Psi_{occ}^\epsilon + \Psi_{unocc}^\epsilon$ , where  $\epsilon$  is an eigenvalue of  $\hat{C}_{m_1 m_2}$ . Suppose that in the  $\epsilon$ -sector, all eigenstates are generated by  $\psi_{\epsilon, q}^\dagger$ , with  $q = 1, \dots, N * N_{orb}/(m_1 m_2)$ . The operators  $\psi_{\epsilon, q}^\dagger$ 's thus satisfy (a)  $\hat{C}_{m_1 m_2} \psi_{\epsilon, q}^\dagger \hat{C}_{m_1 m_2}^\dagger = \epsilon \psi_{\epsilon, q}^\dagger$  and (b)  $\{\psi_{\epsilon, q}, \psi_{\epsilon', q'}^\dagger\} = \delta_{\epsilon\epsilon'} \delta_{qq'}$ .

The previously defined operators  $f_{\epsilon_k, \alpha}(\mathbf{r}_i)$  can be decomposed into a linear superposition of all single particle eigenstates:

$$f_{\epsilon_k, \alpha}(\mathbf{r}_i) = \sum_{\lambda} W_{i\alpha, q}^{(\lambda)} \psi_{\lambda, q}, \quad \mathbf{r}_i \in A_0. \quad (C4)$$

But according to the property (a) of  $f_{\epsilon_k, \alpha}(\mathbf{r}_i)$  and property (a) of  $\psi_{\lambda, q}$ , we have  $W^{(\lambda)} \neq 0$  only if  $\lambda = \epsilon_k$ , and therefore

$$f_{\epsilon_k, \alpha}(\mathbf{r}_i) = W_{i\alpha, q}^{(\epsilon_k)} \psi_{\epsilon_k, q}. \quad (C5)$$

One is reminded that  $\mathbf{r}_i \in A_0$ , which is  $1/m_2$  of  $A$  and is therefore  $1/(m_1 m_2)$  of the whole lattice, therefore  $W^{(\epsilon)}$  is in general a  $N * N_{orb}/(m_1 m_2)$ -dimensional unitary matrix. Before proceeding, one is reminded of the difference between the matrix  $U$  introduced in Sec.II(A), which is a unitary transform between the *unsymmetrized*  $c_{\alpha}(\mathbf{r})$ 's and the *symmetrized* single particle eigenstate operators, and the matrix  $W$ , which is a transform between the *symmetrized* operators  $f_{\epsilon, \alpha}$  and the *symmetrized* single particle eigenstate operators. In general,  $W$  is a block-diagonal matrix with each block labeled by its  $\lambda$ . Each block in a block-diagonalized matrix must be unitary, if the whole matrix is unitary.

Using Eq.(C5), one can calculate  $D_k$ :

$$\begin{aligned} (D_k)_{i\alpha, j\beta} &= \sum_{q, q'} W_{q, i\alpha}^{(\epsilon_k)\dagger} W_{j\beta, q'}^{(\epsilon_k)} \langle \psi_{\epsilon_k, q}^\dagger \psi_{\epsilon_k, q'} \rangle \\ &= \sum_{q \in occ} W_{q, i\alpha}^{(\epsilon_k)\dagger} W_{j\beta, q}^{(\epsilon_k)}. \end{aligned} \quad (C6)$$



And  $D_k^2$ :

$$\begin{aligned} (D_k^2)_{i\alpha,j\beta} &= \sum_{m,\gamma} (D_k)_{i\alpha,m\gamma} (D_k)_{m\gamma,j\beta} \\ &= \sum_{q,q' \in occ} W^{(\epsilon_k)^\dagger}_{q,i\alpha} W^{(\epsilon_k)}_{m\gamma,q} W^{(\epsilon_k)^\dagger}_{q',m\gamma} W^{(\epsilon_k)}_{j\beta,q'}. \end{aligned} \quad (C7)$$

Using the unitarity of  $W^{(\epsilon)}$  we have

$$W^{(\epsilon_k)}_{m\gamma,q} W^{(\epsilon_k)^\dagger}_{q',m\gamma} = (W^{(\epsilon_k)^\dagger} W^{(\epsilon_k)})_{qq'} = \delta_{qq'}. \quad (C8)$$

Substituting Eq.(C8) into Eq.(C7), we have

$$(D_k^2)_{i\alpha,j\beta} = \sum_{q \in occ} W^{(\epsilon_k)^\dagger}_{q,i\alpha} W^{(\epsilon_k)}_{j\beta,q} = (D_k)_{i\alpha,j\beta}. \quad (C9)$$

From the explicit expression of Eq.(C6), it is straightforward to see that for every  $q \in occ$ , the column vector  $W^{(\epsilon_k)*}_{i\alpha,q}$  is an eigenvector of  $D_k$  of eigenvalue 1 and for every  $q \notin occ$ , it is an eigenvector of eigenvalue zero. Therefore  $\dim(D_k^{(s)})$  is equal to the number of occupied states with  $C_{m_1 m_2}$  eigenvalue  $\epsilon_k$ , or the number of occupied states with  $C_n$  eigenvalue  $\lambda$  that satisfies  $\lambda^{\frac{n}{m_1 m_2}} = \epsilon_k$ .

#### Appendix D: Number of in-gap eigenvalues of a matrix $C = \frac{1}{n}(D_1 + \dots + D_n)$ , where $D_i$ is a projector

In this Appendix, we prove that if a matrix  $C$  can be written as the average of  $n$  projectors, i.e.,  $C = 1/n(D_1 + \dots + D_n)$ , then there must be at least  $N_{mid}$  eigenstates, the eigenvalues of which,  $\epsilon$ , are in the range

$$\epsilon \in [1/n, 1 - 1/n], \quad (D1)$$

. The number of these states is given by

$$N_{mid} = \max_{i,j=1,\dots,n} |\text{Dim}(D_i) - \text{Dim}(D_j)|, \quad (D2)$$

where  $\text{Dim}(D_i)$  is the number of nonzero eigenvalues of  $D_i$ .

Before going on, the authors would like to point out that the proof is completely mathematical and has very little relation to the physics presented in the paper. It is provided here only for completeness of the work.

First we prove a lemma: Given two hermitian matrices  $A$  and  $B$ , with eigenvalues  $a_i$ 's and  $b_i$ 's, consider their sum  $M = A + B$ , then for any eigenvalue of  $A$ , say  $a_i$ , there must be an eigenvalue of  $M$  that satisfies  $a_i + b_{min} \leq m^* \leq a_i + b_{max}$ , where  $b_{min}$  ( $b_{max}$ ) is the minimum (maximum) eigenvalue of matrix  $B$ . We can always define  $B' = B - (b_{min} + b_{max})/2I$ , while  $B$  and  $B'$  have the same eigenstates, and here the maximum/minimum eigenvalue of  $B'$  is  $\pm(b_{max} - b_{min})/2 \equiv \pm b_0$ . So for any eigenstate of  $M' = A + B'$  whose eigenvalue is between  $a_i - b_0$  and  $a_i + b_0$ , the same state must also be an eigenstate of  $M = A + B$  with eigenvalue between  $a_i + b_{min}$  and  $a_i + b_{max}$ .

Denote the eigenstate of  $A$  with eigenvalue  $a$  with  $|a\rangle$ , then consider the quantity

$$\langle a|B'^2|a\rangle = \sum_n |y_j|^2 b_j'^2 \leq b_0^2, \quad (D3)$$

where we have used the expansion of  $B'$  in terms of its eigenstates  $B' = \sum_j b'_j |b'_j\rangle\langle b_j|$  and the expansion of  $|a\rangle$  in  $|b'_j\rangle$ ,  $|a\rangle = \sum_j y_j |b'_j\rangle$ .

On the other hand, one can always decompose  $|a_i\rangle$  in terms of eigenvectors of  $M'$

$$|a_i\rangle = \sum_n x_n |m'_n\rangle, \quad (D4)$$

where  $|m'_n\rangle$  is an eigenvector of  $M'$  with eigenvalue  $m'_n$ . We again calculate the quantity

$$\begin{aligned} \langle a|B'^2|a\rangle &= \langle a|(M' - A)^2|a\rangle \\ &= \langle a|(M' - a)^2|a\rangle \\ &= \sum_n |x_n|^2 (m'_n - a)^2 \\ &\leq b_0^2. \end{aligned} \quad (D5)$$

Since  $|x_n|^2 \geq 0$  for each  $n$  and  $\sum_n |x_n|^2 = 1$ , there must be at least one  $n^*$  that satisfies

$$|m'_{n^*} - a| \leq b_0, \quad (\text{D6})$$

or

$$a - b_0 \leq m_{n^*} \leq a + b_0. \quad (\text{D7})$$

Therefore for we know there must be an eigenvalue of  $M$ ,  $m_{n^*} = m'_{n^*} + (b_{\max} + b_{\min})/2$  that is between  $a_i + b_{\min}$  and  $a_i + b_{\max}$ .

Consider the matrix  $D = D_1 + \dots + D_n$ , define  $A = D_1 + D_2$  and  $B = D_3 + \dots + D_n$ . If  $\dim(D_1) \neq \dim(D_2)$ , then there must be at least  $\delta_{12} = |\dim(D_1) - \dim(D_2)|$  number of eigenstates of  $A$  with eigenvalue 1. On the other hand, since each  $D_i$  is a projector,  $b_{\min} = 0$  and  $b_{\max} = n - 2$ . Suppose  $|\phi_{1,\dots,\delta_{12}}\rangle$  are the degenerate eigenstates of  $A$  with eigenvalue 1, and they make a subspace called  $\Phi_A$ . Take  $|a_1\rangle = |\phi_1\rangle$ , from the lemma, we can prove that there must be one eigenstate  $|\lambda_1\rangle$  with eigenvalue between  $[1, n - 1]$ . Then if  $\delta_{12} > 1$ , one is guaranteed to have a state within  $\Psi_A$  that is orthogonal to  $|\lambda_1\rangle$ . Take that state as  $|a_2\rangle$ . The decomposition of  $|a_2\rangle$  does not contain  $|\lambda_1\rangle$ , or,  $x_{\lambda_1} = 0$ . Therefore, for Eq.(D5) to hold, there must be another eigenstate of  $M$ , called  $|\lambda_2\rangle$  that has eigenvalue between  $[1, n - 1]$ . Then if  $\delta_{12} > 2$ , one is guaranteed to have a state within  $\Phi_A$  that is orthogonal to both  $|\lambda_{1,2}\rangle$ , called  $|a_3\rangle$ . Start from  $|a_3\rangle$ , one obtains another eigenstate of  $M$  with eigenvalue between  $[1, n - 1]$ . The process continues until one has exactly  $\delta_{12}$  eigenstates of  $M$  with eigenvalues  $[1, n - 1]$ . Now, one is *not* guaranteed to have a state in  $\Phi_A$  that is orthogonal to all  $|\lambda_{1,\dots,\delta_{12}}\rangle$ . In this way, we have proved that there must be at least  $|\dim(D_1) - \dim(D_2)|$  eigenstates of  $D$  having eigenvalues between 1 and  $n - 1$ . It means there are  $|\dim(D_1) - \dim(D_2)|$  number of protected in-gap state in the spectrum of  $C = D/n$ .

Notice that in the proof,  $D_1$  and  $D_2$  are arbitrarily chosen, and we can in principle choose any two  $D_i$ 's. This means the least number of protected in-gap states is

$$N_{\text{mid}} = \max_{i,j=1,\dots,n} |\dim(D_i) - \dim(D_j)|, \quad (\text{D8})$$

and these states are protected because their eigenvalues must stay within the range  $[1/n, 1 - 1/n]$  and can never approach 0 or 1 infinitely.

### Appendix E: Explicit derivation of the number of in-gap states for a $A_{1/2}^3$ -cut in a $C_6$ -invariant system

In this Appendix we explicitly show how we derive the expression of the number of protected in-gap states in terms of the  $Z^n$  index for a  $A_{1/2}^3$  cut in a  $C_6$ -invariant insulator.

In this system, the index  $z_{m=1,\dots,6}$  represents the number of occupied states in the  $m$ -th representation of  $C_6$ , that is, the states that have eigenvalue  $\exp(i\pi(F + 2m - 2)/n)$  under  $C_6$ . According to the process sketched in Sec.IV, first we can block diagonalized the correlation matrix  $C(A_{1/2}^3)$  into three blocks. Then the total number of in-gap states is simply the sum of the numbers of in-gap states in each block. In the first block describes the entanglement between states with  $C_3$  eigenvalue  $\exp(iF\pi/3)$  inside and outside  $A$ ; the second block describes the entanglement between states with  $C_3$  eigenvalue  $\exp(i(F + 2)\pi/3)$  inside and outside  $A$ ; the third block describes the entanglement between states with  $C_3$  eigenvalue  $\exp(i(F + 4)\pi/3)$  inside and outside  $A$ .

Next, we notice that states with the same  $C_3$ -eigenvalue, i.e., contributing to the same block, in general have two different  $C_6$ -eigenvalues. States have  $C_3$ -eigenvalue  $\exp(iF\pi/3)$  may have  $C_6$ -eigenvalue  $\exp(iF\pi/6)$  or  $\exp(i(F + 6)\pi/6)$ ; states with  $C_3$ -eigenvalue  $\exp(i(F + 2)\pi/3)$  may have  $C_6$ -eigenvalue  $\exp(i(F + 2)\pi/6)$  or  $\exp(i(F + 8)\pi/6)$ ; states with  $C_3$ -eigenvalue  $\exp(i(F + 4)\pi/3)$  may have  $C_6$ -eigenvalue  $\exp(i(F + 4)\pi/6)$  or  $\exp(i(F + 10)\pi/6)$ . Therefore, an occupied subspace with a certain  $C_3$ -eigenvalue can be split into two subspaces with different  $C_6$ -eigenvalues. The number of in-gap states in each block is exactly equal to the difference between the number of states in these two  $C_6$ -subspaces.

According to this analysis, we have  $N_{\text{mid}}^{(1)} = |z_1 - z_4|$ ,  $N_{\text{mid}}^{(2)} = |z_2 - z_5|$  and  $N_{\text{mid}}^{(3)} = |z_3 - z_6|$ . The total number of in-gap states is  $N_{\text{mid}} = N_{\text{mid}}^{(1)} + N_{\text{mid}}^{(2)} + N_{\text{mid}}^{(3)} = |z_1 - z_4| + |z_2 - z_5| + |z_3 - z_6|$ .

### Appendix F: Explicit derivation of $\text{Dim}(\Psi_{\text{occ}}^{e^{i(F+4)\pi/6}})$ in Table II

In this Appendix, we use the example of counting the dimension of  $\Psi_{\text{occ}}^{e^{i(F+4)\pi/6}}$  to illustrate the way we use to derive Table II.

From Appendix C, we know that  $\text{Dim}(\Psi_{occ}^{e^{i(F+4)\pi/6}})$  is the number of occupied states whose  $C_6$ -eigenvalue is  $e^{i(F+4)\pi/6}$ . We have also shown that only high symmetry points contribute to the *difference* of the dimensions of subspaces, and therefore we ignore the contribution from states at generic  $\mathbf{k}$ 's.

At  $\Gamma$ , the highest symmetry is  $C_6$ , so the contribution to  $\text{Dim}(\Psi_{occ}^{e^{i(F+4)\pi/6}})$  is simply the number of states with  $C_6$  eigenvalue  $e^{i(F+4)\pi/6}$ . At  $K$ , the highest symmetry is  $C_3$ . Notice that  $C_3 = C_6^2$ , and we can see that if the  $C_3$ -eigenvalue of a state at  $K$  is  $e^{i(F+4)\pi/3}$ , then it will contribute +1 to  $\text{Dim}(\Psi_{occ}^{e^{i(F+4)\pi/6}})$ . Finally,  $M$  is a  $C_2$ -invariant point. Notice that  $C_2 = C_6^3$ , and we can conclude that a state at  $M$  contributes +1 to  $\text{Dim}(\Psi_{occ}^{e^{i(F+4)\pi/6}})$  if and only if its  $C_2$ -eigenvalue is  $\exp(iF\pi/2)$ .

- 
- <sup>1</sup> X.L. Qi, Y.S. Wu, and S.C. Zhang, Phys. Rev. B **74**, 085308 (2006).
  - <sup>2</sup> C. L. Kane and E. J. Mele, Physical Review Letters **95** (2005).
  - <sup>3</sup> B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).
  - <sup>4</sup> L. Fu and C. L. Kane, Physical Review B (Condensed Matter and Materials Physics) **74**, 041405 (2006).
  - <sup>5</sup> J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007).
  - <sup>6</sup> M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
  - <sup>7</sup> L. Fu, C. L. Kane, and E. J. Mele, Physical Review Letters **98**, 106803 (2007).
  - <sup>8</sup> D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature **452**, 970 (2008).
  - <sup>9</sup> D. Hsieh, Y. Xia, D. Qian, L. Wray, J. H. Dil, F. Meier, L. Patthey, J. Osterwalder, A. Fedorov, A. B. H. Lin, D. Grauer, Y. Hor, R. Cava, and M. Hasan, Nature **460**, 1101 (2009).
  - <sup>10</sup> H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Nat. Phys. **5**, 438 (2009).
  - <sup>11</sup> Q. Liu, C.-X. Liu, C. Xu, X.-L. Qi, and S.-C. Zhang, Phys. Rev. Lett. **102**, 156603 (2009).
  - <sup>12</sup> X. Zhou, C. Fang, W.-F. Tsai, and J. Hu, Phys. Rev. B **80**, 245317 (2009).
  - <sup>13</sup> H. Yao and S. A. Kivelson, Phys. Rev. Lett. **105**, 166402 (2010).
  - <sup>14</sup> X.-L. Qi, T. L. Hughes, and S.-C. Zhang, (2010), arXiv:1003.5448.
  - <sup>15</sup> X. L. Qi and S. C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).
  - <sup>16</sup> K. Sun, W. V. Liu, A. Hemmerich, and S. Das Sarma, Nature Physics **8**, 67 (2011).
  - <sup>17</sup> X. Wan, A. Turner, A. Vishwanath, and S. Y. Savrasov, Phys. Rev. B **83**, 205101 (2011).
  - <sup>18</sup> G. Xu, H. Weng, Z. Wang, X. Dai, and Z. Fang, Phys. Rev. Lett. **107**, 186806 (2011).
  - <sup>19</sup> C. Fang, M. J. Gilbert, X. Dai, and B. A. Bernevig, Phys. Rev. Lett. **108**, 266802 (2012).
  - <sup>20</sup> G. B. Halász and L. Balents, Phys. Rev. B **85**, 035103 (2012).
  - <sup>21</sup> K. v. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).
  - <sup>22</sup> D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Physical Review Letters **49** (1982).
  - <sup>23</sup> F. D. M. Haldane, Physical Review Letters **61** (1988).
  - <sup>24</sup> X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Physical Review B (Condensed Matter and Materials Physics) **78**, 195424 (2008).
  - <sup>25</sup> M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
  - <sup>26</sup> A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, Phys. Rev. B **78**, 195125 (2008).
  - <sup>27</sup> S. Ryu, J. E. Moore, and A. W. W. Ludwig, arxiv: 1010.0936.
  - <sup>28</sup> M. Nakahara, *Geometry, Topology, and Physics* (A. Hilger, 1990).
  - <sup>29</sup> X.L. Qi, Y.S. Wu, and S.C. Zhang, Phys. Rev. B **74**, 045125 (2006).
  - <sup>30</sup> A. Chandran, M. Hermanns, N. Regnault, and B. A. Bernevig, e-print arXiv: 1102.2218 (2011).
  - <sup>31</sup> L. Fu, Phys. Rev. Lett. **106**, 106802 (2011).
  - <sup>32</sup> T. L. Hughes, E. Prodan, and B. A. Bernevig, Phys. Rev. B **83**, 245132 (2010).
  - <sup>33</sup> A. Alexandradinata, X. Dai, and B. A. Bernevig, unpublished.
  - <sup>34</sup> A. M. Turner, Y. Zhang, R. S. K. Mong, and A. Vishwanath, Phys. Rev. B **85**, 165120 (2012).
  - <sup>35</sup> H. Li and F. D. M. Haldane, Phys. Rev. Lett. **101**, 010504 (2008).
  - <sup>36</sup> R. Thomale, D. P. Arovas, and B. A. Bernevig, Phys. Rev. Lett. **105**, 116805 (2010).
  - <sup>37</sup> E. Prodan, T. L. Hughes, and B. A. Bernevig, Phys. Rev. Lett. **105**, 115501 (2010).
  - <sup>38</sup> A. Turner, Y. Zhang, and A. Vishwanath, Phys. Rev. B **82**, 241102 (2010).
  - <sup>39</sup> F. Pollmann, A. M. Turner, E. Berg, and M. Oshikawa, Phys. Rev. B **81**, 064439 (2010).
  - <sup>40</sup> L. Fidkowski, Phys. Rev. Lett. **104**, 130502 (2010).
  - <sup>41</sup> I. Peschel, J. Phys. A: Math. Gen. **36**, L205 (1993).
  - <sup>42</sup> A. Alexandradinata, T. L. Hughes, and B. A. Bernevig, Phys. Rev. B **84**, 195103 (2011).
  - <sup>43</sup> X. Chen, Z.-C. Gu, and X.-G. Wen, Phys. Rev. B **82**, 155138 (2010).
  - <sup>44</sup> A. Kitaev and J. Preskill, Phys. Rev. Lett. **96**, 110404 (2006).
  - <sup>45</sup> S. Ryu and Y. Hatsugai, Phys. Rev. B **93**, 245115 (2006).
  - <sup>46</sup> R. Thomale, A. Sterdyniak, N. Regnault, and B. A. Bernevig, Phys. Rev. Lett. **104**, 180502 (2010).
  - <sup>47</sup> A. M. Turner, F. Pollmann, and E. Berg, Phys. Rev. B **83**, 075102 (2011).
  - <sup>48</sup> L. Fidkowski and A. Kitaev, Phys. Rev. B **83**, 075103 (2011).
  - <sup>49</sup> A. Rosch, arXiv:1203.5541v1 (2012).